INTUITIONISTIC SEMANTICS AND THE REVISION OF LOGIC

Bernhard Weiss

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews

1992

INTUITIONISTIC SEMANTICS AND THE REVISION OF LOGIC

by

Bernhard Weiss

A thesis submitted to the faculty of arts of the University of St Andrews in fulfilment of the requirements for the degree of Ph.D.

October 1991

Th B 101

CONTENTS

Declarations

I, Bernhard Weiss, hereby certify that this thesis, which is approximately 80 000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

date .... signature of candidate ....

I was admitted as a research student under ordinance No. 12 in October 1986 and as a candidate for the degree of Ph.D. in June 1988; the higher study for which this is a record was carried out in the University of St. Andrews between 1986 and 1990.

date .... signature of candidate ....

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Ph.D. in the University of St. Andrews and that the candidate is qualified to submit this thesis in application for that degree.

date .... signature of supervisor ....

date .... signature of supervisor ....

Abstract

In this thesis I investigate the implications, for one's account of mathematics, of holding an anti-realist view. The primary aim is to appraise the scope of revision imposed by anti-realism on classical inferential practice in mathematics. That appraisal has consequences both for our understanding of the nature of mathematics and for our attitude towards anti-realism itself. If an anti-realist position seems inevitably to be absurdly revisionary then we have grounds for suspecting the coherence of arguments canvassed in favour of anti-realism. I attempt to defend the anti-realist position by arguing, i) that it is not internally incoherent for anti-realism to be a potentially revisionary position, and ii) that an anti-realist position can, plausibly, be seen to result in a stable intuitionistic position with regard to the logic it condones. The use of impredicative methods in classical mathematics is a site of traditional intuitionistic attacks. I undertake an examination of what the anti-realist attitude towards such methods should be. This question is of interest both because such methods are deeply implicated in classical mathematical theory of analysis and because intuitionistic semantic theories make use of impredicative methods. I attempt to construct the outlines of a set theory which is anti-realistically acceptable but which, although having no antecedent repugnance for impredicative methods as such, appears to be too weak to offer an anti-realistic vindication of impredicative methods in general. I attempt to exonerate intuitionistic semantic theories in their use of impredicative methods by showing that a partial order relying on the nature of our grasp of the intuitionistic meaning stipulations for the logical constants precludes a possible circularity.

Acknowledgements

During my time at St. Andrews I have had the benefit of being taught by a number of people, all of whom deserve my thanks. Stephen Read tutored me for my first term here and was my first ever tutor in philosophy. His thorough reading of my essays and, in particular, the stress he laid on presenting one's thought (or muddle) clearly were a good introduction to the difficult business of thinking and writing about philosophy.

When I look back to my first year in St. Andrews I am aware of just how lucky I am to have had the opportunity, so soon after having begun to study philosophy, of talking regularly with Crispin Wright. I remember that period as having been a period of intense intellectual excitment. That excitment has produced a lasting enthusiasm in me for the subject. Of course I am convinced that that enthusiasm is provoked by something intrinsic to the subject but I must thank Crispin Wright for having communicated that something so thoroughly.

For this thesis I have been supervised by both Peter Clark and Bob Hale. Peter has been around from the begining of the thesis to its end. I am grateful to him for having been so encouraging about the former and for those reassuring sessions when I left his office with absolute confidence that the latter was going to happen. Peter initially shaped the form of this project but has encouraged its emphasis to be determined by my particular interests, allowing it to become my own.

Bob Hale has flattered me by his immensely thorough reading of my work. His criticisms of my thoughts have always been acute and useful so that the pain of realizing the extent of my muddle and the vagueness and obscurity of my writing has always been allayed by knowing that my efforts were being taken seriously. Bob has high standards of lucidity and a huge appetite for philosophy so is an exhausting person with whom to have a supervision session, but that tiredness has always been accompanied by some feeling of progress.

At various stages of my years in the Department of Logic and Metaphysics there has been a valuable body of graduate students. Two fellow students deserve special mention here. Michael Robertson, by his focussed resistance to the sort of theorising which I have gone in for and through his extremely articulate nausea about using terms such as "canonical", has made me think far more deeply than I otherwise might have done about the plausibility of my approach and its relation to more general philosophical concerns. To Marguerite Nesling I owe an unquantifiable debt. Her thinking about this subject has always struck me as penetrating, clear and exciting. In many chats with her I have been impressed both by the substance of her views and by her utterly genuine and irrepressible desire simply to understand.

I should also like to thank my examiner, Dr. Timothy Williamson, whose comments made apparent a number of gaps and errors in the first version of this thesis.

This, in lieu of a grandchild, is for my parents.

... (for only Nobody knows
where truth grows why
birds fly and
especially who the moon is.
(Cummings 1931, LVIX)

... will you come with me into
the extremely little house of
my mind. ...
(Cummings 1931, LVI)

Introduction:

> ...the objection to *impredicativity*, which is the intuitionist ground for rejecting much of classical set theory, has little or no connection with the insistence upon verificationism itself. Indeed intuitionistic mathematics is itself "impredicative", in as much as the intuitionist notion of constructive proof presupposes constructive proofs which refer to the totality of *all* constructive proofs. (Putnam 1983, p.21)

This tantalizing paragraph, occurring towards the end of Hilary Putnam's splendid "Models and Reality", suggests a programme in the philosophy of mathematics. The promise of the programme is a final position which vindicates most of the methods of classical mathematics yet which takes full account of the modern semantic arguments for intuitionism propounded most forcefully by Michael Dummett.

What is the nature of traditional intuitionistic and constructive objections to the use of impredicative methods in classical truth conditional settings?; How do these objections transfer to a non-classical approach?; Or, crudely, if we change our perception of the practice (to one given by an intuitionistic semantics) do we still need to revise the practice itself?; Where do traditional intuitionistic accounts have use for impredicative methods?; Is this use inevitable?; Is it justifiable?

These important questions about the coherence and nature of our programme are among those I have in mind in much of what follows. But the form of my discussion is not moulded by these questions. Instead I adopt a Dummettian framework: I elucidate, briefly, Dummett's "negative" argument against and his characterization of realism and then attempt an appraisal of the positive programme of giving an account of mathematics. The fecundity of Dummett's philosophy has two aspects, the first is his powerful and general argument against realism but, almost of greater importance, is the framework established in the course of this argument. Dummett succeeds in giving us a precise clarification of the nature of the metaphysical dispute about realism by showing how a metaphysical disagreement can be traced back to a divergence in the notion of truth accepted by each party to the dispute, how the notion of truth is linked to a model of understanding and how general semantic considerations provide principles for deciding such disputes. I endeavour to exploit the insight given us by Dummett, so where some of the questions of the previous paragraph surface the Dummettian framework should help clarify the nature of the issue being raised. For instance, it is often unclear what form of justification of a practice is required or of how we are to make use of a metaphysical picture. So objections to or arguments in support of impredicativity become difficult to appraise from any perspective other than that given by a paricular use of a particular metaphysical picture. The region of disagreement threatens to become hopelessly vague. Because Dummett concentrates first on delineating an area of mutual understanding the appearance of controversy is clearly marked.

However the coherence of Dummett's philosophy depends on the possibility of carrying out his positive programme, i.e., if it is not possible to describe a practice in a manner that satisfies Dummettian adequacy constraints on a semantic account then rather than question the practice as a whole we should wonder whether or not those constraints are acceptable. So I want both to use the Dummettian framework as far as possible to illuminate aspects of mathematical practice and, in the light of its ability or inability to acount for the practice, to reflect on the coherence of Dummett's semantic position. So the implications and feasibility of this positive programme form the central and motivating issues of this work.

Had the emphasis of my interests been less closely linked to assesing the viability and consequences of anti-realism it would have been possible to have chosen a more historical approach. The basic question simply concerns which set of inferential techniques we take to be justified in mathematics. One way of looking at that problem is to consider whether and why we are forced to revise classical inferential principles. Historical and methodological reasons combine to make this an inviting strategy. The methodological reason is that we are concerned with possible ways of motivating or of criticizing a realist outlook. Below I give an account of Dummett's characterization of realism and his reasons for thinking that we need to depart from that perspective. I have thus stressed a particular methodological approach to the question of justification and revision of inference.

But historically it has also been true that realist views have, at least impicitly, dominated the development of mathematics. Those realist presuppositions were challenged by the emergence of contradictions in the foundational edifices whose construction was attempted in the second half of the last century. Mathematicians involved in the foundational effort were, to be sure, often motivated by or led towards a particular philosophical or metaphysical view of mathematics. But that view, because of the basic agreement about acceptable mathematical methods, could largely be ignored when appraising the significance of the mathematical work itself. Whatever philosophical motivation Weierstrass may have entertained is largely irrelevant to his work in making the calculus rigorous. Dedekind's view that real numbers are free creations of human thought can be and has been ignored by mathematicians utilising his means of "cuts" for defining natural numbers. Cantor's theological perspective not only failed to impress mathematicians but was never an integral part of understanding Cantorian set theory and only became a feature of Cantor's intellectual life after his discovery of his paradox (and thus his need for a notion of the absolutely infinite). The discovery of the paradoxes led to a questioning of the justification of mathematical methods themselves (rather than the justification of particular mathematical methods in terms of more generally accepted theory). Rather than being interested in the metaphysical attitudes of mathematicians we are concerned with the metaphysical commitments underlying their use of certain techniques.

Thus, historically, philosophy became deeply and explicitly implicated in the foundations of mathematics with the need to depart, in practice, from an out and out realism about mathematics. But let us be a little clearer about the context in which the paradoxes emerged. The problem mathematicians had been preoccupied with was bringing rigour to mathematical analysis. The paramount concern here was to establish analysis without relying on geometrical intuition or without relying on the dubious notion of infinitesimals. This arithmetization of analysis led to Dedekind's method of defining the reals in terms of divisions of the rationals into two disjoint classes, where any number of the one class is less than any member of the other. It also led to Cantor's development of transfinite arithmetic and, in this theory, to the set concept as being fundamental for mathematics. Frege's work, whilst being sympathetic to many of the technical results achieved by Cantor, differs fundamentally in not taking sets as basic aggregates but rather considers classes as extensions of predicates. Essential to this move was Frege's development of the logic of quantification. Thus, for Frege the foundational problem changes since, rather than involving a reduction of one region of mathematics to another realm of essentially mathematical entities, Frege sees that arithmetic should be reducible to purely logical concepts, i.e., to concepts which can be defined purely by logical means.

It is surely true that philosophy became deeply involved in Frege's logicist foundational enterprise and thus does pre-date the emergence of the paradoxes. This is evidenced by Frege's systematic development of philosophical principles which are intended to guide both the analysis of mathematics and of language quite generally. So Frege's programme makes as explicit as possible some of the realist assumptions of classical mathematics and attempts to justify them systematically. It is for this reason that the emergence of the paradoxes in his system are so philosophically compelling. They demand some revision in realist views hitherto tacitly accepted in accepting classical logical practice. So, one of the lessons that Dummett draws from Frege is a characterization of realism.

The paradoxes, or, rather Russell's attempt to deal with the paradoxes, thus provide an apposite point of departure for this discussion. Since my interest here is more narrowly confined to surveying some of the consequences and the coherence of the anti-realist view I have not attempted a systematic catalogue of attempts to account for and solve the paradoxes. Instead I have drawn from my discussion of Russell a scepticism about any possible foundational rôle for logic and have used that to motivate a Dummettian description of the relation between logic and a theory of meaning.

My use of the terms "anti-realist" and "intuitionist" requires some comment. The former is always used to refer to a semantic position which accepts certain epistemic constraints on the central notion of the semantic theory, or, better, it is any position which challenges a realist semantics for using a notion which is not explained appropriately in terms of our recognitional capacities. I use the second term far more loosely. On occasion it is to be taken as synonomous with anti-realism (as applied to mathematics) but more generally it is taken to be a position which adheres to some version of intuitionistic logic and thus is a position marked out simply by the extent of its departure from classical logic, rather than by its motivation for that departure. I hope that context provides sufficient disambiguation. The only point, though, at which the distinction is crucial is in my discussion of strict finitism where the question is whether *anti-realism* collapses into a strict finitist position or whether it motivates an intuitionist position.

CHAPTER ONE: RUSSELL: FROM *THE PRINCIPLES OF MATHEMATICS* TO *PRINCIPIA MATHEMATICA*

Russell: From *The Principles of Mathematics* to *Principia Mathematica:*

The main focus of my discussion in this chapter is the development of Russell's thought from *The Principles of Mathematics* to *Principia Mathematica*. I hope, though, to show that the interest of this investigation is not confined purely to the biography of Russell.

I assume that the classical theory of classes, the (so-called) naive theory of classes betokens, what Michael Dummett would call, a naive realism with respect to classes. Since that theory leads to contradiction naive realism about classes must be wrong. That is, the paradoxes demand "some change in current logical assumptions" (to use Russell's phrase). The interesting question then seems to be: How is this revision in logical practice to be motivated?

Now anyone familiar with Dummett's work will be aware of his forceful arguments for the view that questions of the soundness of a logical system should be appraised from within the theory of meaning. In particular in *The Philosophical Basis of Intuitionistic Logic* he argues that an attempt to motivate a departure from classical logic in favour of intuitionistic logic that is *not* based on a view about meaning but *is* based on a specific ontological view of the status of mathematical objects must either fail to motivate any revision or must be based on a severe scepticism (in this case about the truth values of subjunctive conditionals).

The promise of the meaning theoretic approach is thus, first that it offers a well-motivated reason for rejecting realism and, secondly, that the revision imposed on our first order practice is, arguably, not too radical. (Assuming, of course, that you have no antecedent repugnance for *any* philosophically inspired revision.)

What I think we find in Russell is an attempt to motivate a revision in logical practice which is *not* based on meaning-theoretic arguments. His work thus offers a possible counterexample to Dummett's broad position. Since Russellians are now a fairly rare philosophical species, it would seem that this is not a counterexample worth taking very seriously. What I try to demonstrate here is that diagnosing the lacuna in Russell's programme *is* instructive in that it offers some support for the Dummettian programme and indicates how a Dummettian might deal with some of the problems to do with impredicativity.

§1 Constraints on a Solution to the Paradoxes:

Russell (1906a) after having examined a selection of the set theoretic paradoxes concludes,

> ... that all of them belong to a certain type, and that none of them are essentially arithmetical, but all belong to logic, and are to be solved, therefore, by some change in current logical assumptions. (p.144)

The reason that Russell gives for holding that the paradoxes point to a failing in "current logical assumptions" is that all the known contradictions can be manufactured according to a "recipe" for constructing what Russell calls "self-reproductive" classes and processes. That is, current logical assumptions permit properties which satisfy the following,

> Given a property φ and a function f, such that, if φ belongs to all the members of u, $f'u$ always exists, has the property φ, and is not a member of u; then the supposition that there is a class w of all terms having the property φ and that $f'w$ exists leads to the conclusion that $f'w$ both has and has not the property φ. (1906c, p.199)

e.g., take φ to be non-self-membership, *f* to be the identity map. Then if all the members of a class *u* have the property φ, i.e., are non-self-membered, then, in particular, *u* is not a member of *u* (since if it were a member of itself we would contradict the assumption that all members of *u* have the property φ, i.e., are non-self-membered). So *f'u* (i.e., *u*) exists and has the property φ. *w* is then the paradoxical class of non-self-membered classes. Since, on the one hand, we can substitute *w* for *u* and deduce that *w* is not a member of *w*. Whilst, conversely, since φ, i.e., non-self-membership belongs to all members of *w*, we know that *w* should have the property φ and thus is a member of *w*. For the Burali-Forti contradiction take φ to be the property of being a well-ordered series of ordinals, *f* to be the map from a well-ordered series to its ordinal number, then if *u* is a series of ordinals the map *f* from *u* to its ordinal takes *u* to a set having the property of being a well ordered series of ordinals which is not included in *u* (since the ordinal of a series of ordinals cannot be included in that series). *w* will be the series of all ordinals. *f'w* then both is and is not a series of ordinals because, on the one hand, *w* is the series of *all* ordinals so *f'w* must be included in *w*. But, on the other hand, since *f'w* is the ordinal of *w* it must occur in the series after *w*, so cannot be included in *w*.

Russell takes this specification of the general form of the contradictions to be important because he regards it as showing that all the contradictions arise from certain logical assumptions; first, about which properties exist and, second, that a property always determines a class. Thus the contradictions are to be treated as a challenge to specifically logical assumptions: arithmetic assumptions about the nature of the transfinite are not implicated in the above specification. In (1906c) Russell urges that Poincaré is wrong in laying the blame for the paradoxes on the assumption that there is an actual infinite since, in his view, the semantic paradoxes (which Russell takes to be of a piece with the set theoretic paradoxes) and, in particular, the liar paradox do not involve classes at all (although they can be construed as involving quantification). The problem thus lies in some sort of self-reference.

However Russell, in the same paper, does subscribe to Poincaré's diagnosis of the source of the paradoxes as lying in some kind of vicious circularity. The precise formulation of the Vicious Circle Principle (VCP) need not detain us here. For now, all I want to note is that Russell does not see his task as having been completed on formulating the VCP. The VCP, if implemented as a means of qualifying definitions so as to exempt paradoxical cases will contravene the principle it expresses,

> Having first put E="all numbers definable in a finite number of words", we arrive at a paradox, due, says M. Poincaré, to our having included a number only definable in a finite number of words by means of E. This vicious circle he proposes to avoid by defining E as "all numbers definable in a finite number of words without mentioning E." To the uninitiated, this definition looks more circular than ever. (1906c, pp.196-7)

Russell wants to *solve* the paradoxes. That is, he wants to develop a logical system which is internally constrained in a manner which ensures that the VCP cannot be contravened. Russell thus takes the VCP itself to perform a purely negative role in that it is needed to rule out possible systems.

But Russell does not only require that a solution obeys the VCP and is thus contradiction free. The technical reform we are driven to carry out in the logic itself must be well motivated: in canvassing possible .

objections to the Zig-zag Theory[1] Russell notes that the axioms it calls for "cannot be recommended by any intrinsic plausibility" (1906a, p.147); the Theory of Limitation of Size[2] he finds, contrary to first impressions, to lack plausibility. Similarly, when Russell advocates a version of what has come to be known as his substitutional theory he recommends it not simply for avoiding the contradictions but also for the simplification it brings about in "the fundamental assumptions, the primitive propositions upon which the edifice is built". Russell notes that, "we do not deny that there are such entities [as classes and relations], we merely abstain from affirming that there are" (1906b, p.188).

Thus Russell takes seriously the need to argue for any proposed technical solution from a perspective other than that simply of showing that the proposal manages to ban the emergence of contradictions. Russell wants his solution to have "a certain consonance with common sense". A final constraint placed by Russell on an acceptable solution is that the resulting system should allow the development of an "acceptably" large portion of mathematics. The sorts of consideration which Russell uses and is able to use in constructing such motivational arguments will be the central concern of this chapter. In order to tease out the development of such considerations I now need to canvass some of Russell's early views. In particular, I begin by tracing his early ontological perspective.

---

1. This is the theory that propositional functions "determine clases when they are fairly simple and only fail to do so when they are complicated or recondite." The problem it faces is, as Russell notes, that selection of the axioms which characterize simplicity is guided only by avoidance of the contradictions.
2. This is the theory that all proper classes have a certain size property e.g., "of being capable of being arranged in a well-ordered series ordinally similar to a segment of the series of ordinals in order of magnitude".

§2 Russell's Early Ontological Views;

My starting point is Russell's *Principles of Mathematics* (1903). I want to look at the way in which contradictions of two sorts (the first sort being the reflexive paradoxes mentioned above whilst the second sort concerns plural objects and the distinction between the class *as many* and *as one*) make their appearance within that system and then to go on to show how Russell attempts to use the innovation of the notion of an incomplete symbol, introduced in "On Denoting" (1905), to solve both sorts of contradiction. One of the main tasks of this discussion is to show that Russell implements this programme by placing epistemological constraints on his system of logic. These epistemological constriants, we shall see, are problematic to apply and and fail to give an adequate motivation for Russell's final system.

The ontology of *The Principles* is determined by the following doctrine,

> Whatever may be an object of thought, or may occur in any true or false proposition, or can be counted as *one*, I shall call a *term*. (1903, p.43)

Thus for Russell nothing can be exempted from being a term (or, synonymously, an entity) since to deny that something is a term or an entity would be to make that thing the subject of a proposition. Thus the denial of termhood would presuppose the termhood of the thing itself and would thus be self-contradictory.

Terms come in two varieties; things and concepts. A thing may only occur in a proposition as the subject of the proposition. Concepts may occur in propositions both as the subject of the proposition and as predicates (adjectives) or relating relations (verbs), i.e., as relations in intension. A concept in its nominal form is identical with its verbal or adjectival form. Thus "humanity" and "being human" are both symbols for the same concept, the former gives the nominal form, the latter the adjectival form. The reason for this is that it proves impossible to assert that there is a difference between the two forms. Russell sums this up,

> ... if there were any adjectives which could not be made into substantives without change of meaning, all propositions concerning such adjectives (since they would necessarily turn them into substantives) would be false, and so would the proposition that all such propositions are false, since this in itself turns the adjectives into substantives. But this state of things is self-contradictory. (1903, p.46)

The point is that if we try to assert that *one* as adjective differs from 1 as term then *one* as adjective is being used substantivally. This means that either the proposition is false because when used as a term *one* becomes identical with 1, or else there is a further difference between *one* and 1 over and above the fact that the first is used as an adjective and the second as a term. But this means that all propositions about *one* as adjective are false since such propositions treat *one* substantivally and thus make it a term. So for Russell everything must be capable of appearing as the subject of a proposition. The word "term" has the widest possible scope.

Two difficulties follow from this view since both the unity of the proposition and the existence of plural objects are somewhat mystifying on this picture. The unity of the proposition is problematic because the proposition cannot be treated simply as a list of terms which occur in it. A minimal account will at least have to be sensitive to the fact that some term is not occurring in the proposition substantivally. But this means that there is some (ineffable) difference between the term used substantivally and the term used adjectivally or verbally. The second difficulty concerning plural objects arises because certain predications which hold of a collection, paradigmatically, predication of a specific numerosity, cannot be treated as distributive predications over the members of the collection. For instance, "Horace and Herbert are two" cannot be read as "Horace is two and Herbert is two". This indicates that we need to have a conception of the collection as many (precisely because if we only had a conception of the collection as one we could not say that there was anything other than one object). So we must have a notion of object which includes both terms and essentially plural objects, i.e., the word "object" must have a wider scope than that of "term" (cf. (1903) p.55). This, for Russell, is a matter of some significant degree of logical puzzlement. I shall return to this question when discussing Russell's first theory of types.

I need now to turn to the from of Russell's theory and, in particular, to the account of complex entities. Of paramount importance will be Russell's account of propositions as complex entities. Russell's primary concern is not with language. Indeed he notes that "meaning in the sense in which words have meaning, is irrelevant to logic" (1903, p.47). Russell is interested in cataloguing the manner in which entities combine to form complex entities. A proposition is part of the realm of reference, it is what certain expressions in our language might mean. The proposition is not composed of words but of entities and is itself a complex entity. We can thus see Russell's logical programme as serving ontological ends: logic gives an account of what sort of entities there are and how these may be combined to form complex entities. Logic, by a process of analysis, reveals in terms of the utmost generality what there is.

In addition, Russell wants to analyse judgement on the model of acquaintance, that is, he wants to analyse judgement as a two place relation holding between the judger and the object of judgement, the proposition. Thus when we judge falsely (but meaningfully) there must never-the-less be an object of judgement. False propositions, therefore, must have being, i.e., are complex entities in the same sense that true propositions are complex entities which are possible objects of judgement. There is no ontological distinction between true and false propositions. Rather, truth and falsity are primitive properties of propositions. Even granting the implausibility of there being false propositions this position is patently unsatisfactory since it gives no account of the role of the notion of truth, it is utterly mysterious as to why we should prefer true to false propositions. Russell is far from blind to these problems and they increasingly demand his attention. Thus Russell in 1910 vaunts his Multiple Relation theory of judgement as offering the outlines of an account of judgement (and of propositions) which does distinguish between true and false propositions in terms of whether or not the resulting complex corresponds to a fact. The nub of that theory again consisits in an exploitation of incomplete symbols; propositional expressions are treated as incomplete. Truth is thus finally explained in terms of correspondence. However we've still some distance to travel before we meet that Russell. The Russell of the first few years of the century (certainly until 1907) drew attention to these aspects of his view of judgement (and thus of propositions) but did not accept them as offering a deep challenge to that view: he was inclined to accept that our preference for true propositions is perhaps to be explained as "an ultimate ethical proposition" (1904, p.76).

§3 Denoting Concepts:

We need now to look at Russell's explanation of how propositions can be about entities which are not constituents of the proposition. It is essential that Russell should recognize this possibility for (at least) two reasons. First, it is this feature of propositions which allows thought to transcend the bounds of immediate acquaintance and thus is an important elelment in Russell's anti-idealism. In *The Principles* Russell only hints at this. The thought is far more fully developed in "On Denoting" where it is clearly a consequence of the principle of acquaintance which he there formulates. The Principle of Acquaintance states that,

> in every proposition that we can apprehend (i.e. not only those whose truth or falsehood we can judge of, but in all that we can think about), all the constituents are really entities with which we have immediate acquaintance. (1905, pp 55-6)

Thus if a proposition could never be about something which was not one of its constituents we could not conceive of complexes with which we were not (or could not be) acquainted. This notion of a proposition being "about" something other than a constituent thus gives rise to Russell's distinction between knowledge by acquaintance and knowledge by description: we can have knowledge about certain entities (Russell mentions the centre of mass of the solar system) with which we have no acquaintance. But also we apprehend complexes whose truth value we are unable to determine since they are about entities with which we have no acquaintance. (This anti-idealist position inflates into a full-blown realism when combined with Russell's unargued for assumption that classical logic is valid under its orthodox truth conditional interpretation. Since then all propositions are determinately either true or false.)

Although Russell never explicitly formulates the principle of acquaintance in *The Principles* that work does seem to be tacitly informed by such a view. One reason for holding this is circumstantial, in that Russell does not announce the principle as a change in his view. But more importantly Russell's analysis of propositions is guided (despite his professed lack of interest in language) by grammatical form and considerations about relations between propositions (e.g., such as when I say "I met a man." I am speaking about a particular man with, say, a drunken wife and a pub etc.). When, however he comes to examining what it is to apprehend a proposition (as he does when he examines Meinong's views) he takes assuming a proposition -which is no more than apprehending it- to be equivalent to presenting the proposition to oneself. Presenting a complex is seen on the model of presenting a particular (and to be presented with a particular is just to be acquainted with it). Russell does not commit himself to a particular view of what is involved in presenting a complex. He rejects the idea that presentation of a complex is a complex of presentations since the latter is not guaranteed the unity of the former. However he is also dubious about supposing that a presentation of a complex is to be regarded as a simple presentation. It seems likely that a presentation of a complex demands at least the possibility of the presentation of the constituents. Indeed it is hard to envisage what the process of analysis achieves unless it is linked to some such possibility. This supports the idea that the epistemology implicit in *The Principles* is at least consistent with the principle of acquaintance. Further it is clear from Russell's discussion of infinite classes that he supposes that propositions, or at least propositions which we grasp, can not be infinitely complex. He claims that where we use a concept to denote an infinite class that concept must do so by giving an intensional specification of the class, that, "the concept *all numbers*, though not itself infinitely complex, yet denotes an infinitely complex object" (1903 p.73). "This," he claims, "is the inmost secret of our power to deal with infinity" since an infinitely complex concept "could not be manipulated by the human intelligence" (1903, p.73). If it were not the case that apprehension of a complex demanded some mode of acquaintance with its constituents it would be difficult to see why we could not grasp a complex which was shown, on analysis, to be infinitely complex. The fact that a regress of analysis (as opposed to implication) is anathema to Russell presumably has the same epistemological source. (Although Russell (1918) in response to a question revokes this abhorrence and admits that it is quite possible that analysis should be without end. By that time, however, Russell had adopted the metaphysical position of logical atomism.)

Thus the theory of denoting concepts from its very inception performs something of an epistemological role in accounting for our apprehension of certain complexes. That role is made much more explicit in the account of denoting given in "On Denoting" but that role is never far from Russell's mind, as this passing remark in the opening paragraph of the chapter on denoting in *The Principles* gives witness to,

> This notion [i.e. the logical notion of denoting] lies at the bottom ... of the opposition between things and ideas, discursive thought and immediate perception. (1903, p.53)

The second reason why denoting concepts are important for Russell is that a denoting phrase such as "a dog" occurs in the subject position of a proposition, e.g., "A dog bit her cat.". The sentence is about a specfic dog (as Russell might have said, a specific dog with a kennel in the garden and a grumpy, cat-hating owner) and that dog is, of course, a specific entity. What Russell needs to explain is how an entity corresponding to "a dog" can occur in the subject position of a proposition without that proposition then being about that entity. Denoting concepts have the peculiar property that when they occur in a proposition the proposition is never about the denoting concept itself but about the object denoted by the denoting concept. The relation of denotation obtaining between these two entities always remains somewhat mysterious so it is questionable whether Russell in his use of that notion has done more than label a phenomenon. I shall return to Russell's own criticism of *The Principles* theory of denoting concepts below.

First, we must look a little more closely at the detail of that theory. If *a* is a predicate then any of the following is a denoting concept; all *a*, every *a*, any *a*, an *a*, some *a*, the *a*. Russell's initial discussion of denoting concepts is confined, first, to illustrating the differences between the various denoting concepts and, secondly, to showing that these differences are to be accounted for not in terms of a different manner of denoting in each case but in terms of different objects denoted in each case. The objects denoted will in any given case be a particular combination of terms, the nature of the combination is determined by the particular denoting concept. The combinations so formed are not, or, at any rate, are not always, themselves terms. Thus we have plural and ambiguous objects which are not terms. (I noted above the strain this places on Russell's notion of *term* since it implies that the notion of *object* is wider than the supposedly maximally applicable notion of *term*. Russell thinks that he can live with this anomaly since in any proposition about a plural object, say, "*A* and *B* are 2" there is *no* subject since we can take none of *A*, *B* or *A*-and-*B* to be the subject.)

What sort of objects do the various denoting concepts denote? "All *a*'s" denotes the plural object formed by additively combining all entities which are *a*'s. In other words, "all *a*'s" is the concept of the class of *a*'s (*a* is called the class concept) and the object denoted is the class of *a*'s considered as a class as one. "Every *a*" distributively denotes each term of the class of *a*'s, it denotes the class of *a*'s considered as a class as many. "Any *a*" denotes an ambiguous *a*, i.e., it denotes the permissive disjunction of the terms in the class of *a*'s in the sense that it is irrelevant which *a* we focus on in particular and we may focus on any one. "An *a*" denotes the disjunction of entities which are *a*'s where the disjunction is taken non-permissively as insisting that no one particular *a* must be taken. Finally, "some *a*" denotes the disjunction of entities which are *a* where the disjunction is taken distributively, i.e., where it not irrelevant which *a* is to be considered but the manner of denotation refuses to specify which.

Russell discusses denoting concepts of the form "the *a*" separately. His discussion is brief. Denoting concepts of this form function only when we have a property which is uniquely satisfied. When that is so we can form propositions about the particular object which satisfies the property by means of the definite description. This is useful since many definitions are fertile just because they enable us to focus on one entity whose precise nature is not of importance save for the fact that they uniquely possess a certain property. Statements of identity are shown to be informative when the resulting proposition includes either two definite description denoting concepts (which, of course denote the same entity) or one definite description denoting concept and the entity itself. So we may apprehend propositions containing the denoting concept without being aware of which object is thus denoted. The identity statement either reveals this to us or equates two modes of denotation. Either way we seem to learn something.

§4 Generality, and Necessity:

What I want to go on to discuss is Russell's treatment in *The Principles* of the contradictions. Before doing so and in order to assess how well his proposed solution coheres with other aspects of *The Principles* view I shall need to consider the relation of his views of logic, necessity and generality.

Russell accepts two simple arguments which then guide his construction of the theory of *The Principles*. The first purports to establish the self-contradictory nature of any assertion of non-termhood. The second holds that quantification must always be unrestricted since any condition used to restrict the range of quantification can be incorporated into a conditional statement of the quantified sentence which must then be interpreted as having a greater range (i.e., we can always move from $(\forall x \varepsilon D) F(x)$ to $\forall x(x \varepsilon D . \supset . F(x))$). We are thus led to the view that there is a single universe of (both simple and complex) entities. That simple picture is complicated by having to accept that objecthood has a wider scope than termhood, i.e., that some objects are not terms. Also we have to accept that certain entities stand in a relation of denoting to other entities (and that the entity which denotes is only explained in terms of having this relation with the denoted entity/ies: this means that the nature both of the denoting relation and of the denoting concept are left mysterious).

To appreciate fully the importance for Russell of having a single domain of quantification one needs to have some conception of his idea of necessity and of logic. *The Principles* begins with a definition of mathematics as,

> the class of all propositions of the form "p implies q," where p and q are propositions containing one or more variables, the same in the two propositions, and neither p nor q contains any constants except logical constants. And logical constants are definable in terms of the following: Implication, the relation of a term to a class of which it is a member, the notion of *such that*, the notion of relation and such further notions as may be involved in the general notion of propositions of the above form. In addition to these mathematics *uses* a notion which is not a constituent of the propositions it considers, namely the notion of truth. (1903, p.3)

The natural question raised by this is, what counts as a *logical* constant? Take the following example which most of us would be happy to consider as a logical truth. "If the class of Greeks is contained in the class of men then, x is a Greek implies x is a man, for all values of x." Russell will not count this as a logical truth because, first, it contains the non-logical canstants *Greeks* and *men* and, secondly (and relatedly), the propositions which take the place of p and q do not contain a variable (and thus do not contain a common variable). We arrive at a logical proposition by generalizing on *Greeks* and *men* to obtain, "If the class a is contained in the class b then, x is an a implies x is a b for all values of x, and this holds for all values of a and b." Russell has no conception of a distinction between a proposition being true and it being necessarily true. So our attempt to justify holding the first proposition true by saying that it is true in all possible worlds will be unintelligible for Russell. For Russell,

> there seems to be no true proposition of which there is any sense in saying that it might have been false. One might as well say that redness might have been a taste and not a colour. What is true, is true; what is false, is false; and concerning fundamentals, there is nothing more to be said. (1903, p.454)

He immediately goes on to give his positive view of necessity which he attributes to Moore,

> The only logical meaning of necessity seems to be derived from implication. A proposition is more or less necessary according as the class of propositions for which it is a *premiss* is greater or smaller. (1903, p.454)

In important respects this is just the view of necessity argued for by Moore in *Necessity* (1900). There, Moore argues that necessity is derived from a certain logical relation of priority which one proposition may have to another and that one proposition is more necessary than another if it is logically prior to the second. Some propositions may be absolutely necessary in virtue of having this relation to all other propositions. The relation alluded to is based on implication, although Moore does not give a detailed account of implication itself. Russell's brief account retains important elements of this conception; the relativity of necessity and the linking of the degree of necessity with some notion of generality. For Moore, a proposition A is more necessary than a proposition B if judging B to be true implies (and is not implied by) a judgement that A is true, so e.g., judging that this is black and that is white so this is different from that, implies there are properties of blackness and whiteness and that black is different from white (which is thus logically prior and so more necessary than the first judgement) and ultimately that things with differing properties are different (which is thus more necessary than both previous judgements). Russell's description of this in terms of whether or not one proposition is or is not a premiss of the other in an implication will not take us beyond the view that a proposition if true, is true, and if false, is false. Since, because implication is purely material, whether or not one proposition implies another will depend solely on whether or not it is true. In Russell's terminology, we are not so much concerned with *implies* as with *therefore*, i.e., we are concerned with grounds for holding the truth of a given proposition. Fundamental logical propositions thus are those that are basic in that they have no grounds, logical (and mathematical) propositions have only logical propositions as their grounds and which in turn function as (part of the) the grounds for a large number of other propositions. Logical priority and necessity are thus apt to seem essentially epistemic notions: one proposition is logically prior to another if it is part of the (perhaps ideal) justification of our knowledge of the latter. This perhaps explains the fact that Moore terms absolutely necessary propositions *a priori.* Russell however attempts to draw a contrast between *implies* and *therefore* on non-psychological grounds. The former holds between unasserted propositions and the latter between asserted propositions where the notion of assertion is taken as being logical rather than psychological. But what this mind-independent notion of assertion amounts to is a riddle Russell acknowledges but shunts out of his path by observing that since it is purely logical is not directly relevant to his specific programme of giving an account of mathematics. The problem, of course, is that Russell is supposed to be giving a *logical* account of mathematics so cannot be so cavalier in his treatment of logical problems.

I want to save from this discussion a realization of the depth of the connection between Russell's notions of logic and of generality. Recall our proposition about Greeks and men, Russell would dismiss this as not being logical because it involves the specific nature of *Greek* and of *men.* But his reason for that view is that the proposition is not of ultimate generality: it can be seen as an instance of a more general proposition. (It is hard to see why this view does not defeat Russell's logicism from the outset since the membership relation is a specific instance of relations in general. Thus if mathematics requires this specific relation it cannot be purely logical.) We have now equated logic with the idea that there are propositions of this ultimate level of generality: a logical proposition includes no entities which have not been eliminated in favour of the (unique) variable. If we accept that the paradoxes demand that we acknowledge a hierarchy of different variables (and, as we shall see, the Russell of *The Principles* seems to think that this is so) then we have to have some means of limiting the scope of our logic. The problem then is, given the identification of logic with propositions of ultimate generality, how do we limit the scope of the logic and still maintain that we are doing logic? The problem is not irresolvable for Russell but the position is not congenial to him. Seemingly the solution has to depend upon showing that our putative propositions are of maximal generality in that propositions purporting to have greater generality are in fact non-sensical. Thus Russell is obliged to elucidate the logical relation of meaning.

§5 The Origin of Russell's First Theory of Types:

At this point we need to consider some problems that attend Russell's view of classes. Russell assumes that any concept which is a predicate unproblematically defines a class (*as many*) (1903, p.54). Such concepts always give rise to class-concepts. A concept which denotes a given class is called the concept of the class, this concept may be derived from the class-concept, e.g., *men* is an example of the former and *man* is the associated example of the latter. Men, the object denoted by *men*, is the class itself. But the distinctions do not end here, we still need to distinguish the class *as many* from the class *as one*. The most fruitful way of viewing this distinction is in terms of the different sorts of predication that can be made of classes. In the discussion of denoting I noted that in order to make sense of some predicates we need to distinguish between objects and terms (strictly, between objects that are terms and objects that are not). Objects consist of combinations of terms, we can apply a predicate to the object which does not apply to any of the individual terms (i.e., we can make a non-distributive predication). The most obvious example of this is numerical predication. In making such a predication it is essential that we regard the object as a combination of terms. However we seem also to be able to treat the class as a term in its own right and can assert, for instance, that it is a member of another class. The former sort of predication calls for the concept of the class as many (e.g., men or all men) whilst the latter for the class *as one* (e.g., the class of men).

Allowing the validity of the class *as one* is (as just noted) to treat the class as one as an entity. Two problems (grave ones) arise. First, consider the class paradox of the class of all non-self-membered classes. Here we have a predicate (non-self-membership) which thus defines a class *as many*. If we assume that the corresponding class *as one* exists and (as we have no reason to suspect otherwise) can be included in the class *as many* then we have the paradox (i.e., the class is a member of itself iff it is not). So the paradox shows either that not all classes *as many* are associated with a class *as one* or that the class *as one* must be treated as a distinguished sort of entity. Secondly, Russell sees classes *as many* purely extensionally, i.e., he identifies the class with a cartain numerical conjunction of its members. But if, now, we consider a class α containing more than one term then it must be identical with the class containing only α. So the class α is identical with a class containing only one member. But this contradicts the assumption that α contains more than one member. The conclusion that there are no classes containing more than one member is obviously intolerable.

The solution to this second problem to which Russell is drawn in the main body of *The Principles* involves accepting an ultimate distinction between the class *as many* and the class *as one*. In the above argument we need first to consider α as a class *as many* and then, in considering it as a member of the class containing only α, as a class *as one*. It is with the class α *as one* that, in view of extensionality, we identify the class containing only α. This does not contradict the fact that the class α *as many* contains more than one term.

In the appendix on Frege which Russell wrote later he rejects this solution. Here he repudiates the notion of the class *as one*. This move is in tension with some of the basic tenets of *The Principles*. Russell needs the notion of the class as one because he needs to be able to talk about the class itself, that is, he needs to be able to make the class the subject of a proposition. Anything that can be made a logical subject is a term. If we give up the capacity to make classes logical subjects most of mathematics will simply become impossible, e.g., if we are to give an account of number in terms of classes we shall need to be able to make classes the logical subjects of a proposition. Russell faces this difficulty by reforming his notion of logical subject. Recall that Russell attempts to defuse the impending contradiction of having a logical subject (namely a plural object) which is not a term by arguing that propositions involving plural objects have no logical subject. Russell now accepts that the notion of objecthood forces an ontological distinction to be drawn between possible logical subjects,

> The subject of a proposition may be not a single term, but essentially many terms. ...But the predicates or class-concepts or relations which can occur in propositions having plural subjects are different ... from those that can occur in propositions having single terms as subjects. (1903, p.516)

So our domain of possible logical subjects must be stratified and symbols for predicates (and relations etc.) which seemingly are universally applicable must be taken to stand for a range of predicates (relations etc.) which have definite meanings only relative to logical subjects of a given sort or type. The stratification of logical subjects proceeds simply by considering entities, then classes of entities, then classes of classes of entities etc. .

This thoroughgoing revamping of some of the basic logical doctrines of *The Principles* is intended to solve the basic problem of Russell's earlier view of classes, viz., the contradiction resulting from considering the class of non-self-membered classes. This it obviously achieves in a purely technical sense since the class concept $x \notin x$ is not meaningful; ε is a relation which is nonsense unless asserted to hold between one object and another of type one degree higher than the first (and, plainly, we cannot negate nonsense).

This solution seems to be motivated by the thought that there is something incoherent in our notion of the class *as one* which should thus be eliminated. The effect of the solution is to salvage a great deal of what we find mathematically useful to say about the class *as one*. This is achieved by broadening our notion of logical subject. It is thus hard to read the proposal as one which brings about a slimmer ontology, indeed, quite the reverse seems to be the case. Neither ontological economy nor an epistemological consideration is the primary motivation here. Rather where Russell does try to motivate the view he does so by appealing directly to the supposed logical nature of the thing concerned: the class is *essentially* many and cannot, without lapsing into nonsense, be treated simply as an entity, i.e., as a one. The problem now is that Russell seems to provide us with the means for talking sensibly about the class as one provided, of course, that we realize that *one* as applied to classes differs from *one* as applied to entities. But our grasp of what an entity is was given solely in terms of what could be treated as one. This now must be modified to what can be treated as one *as applied to entities*. And evidently this is circular. Thus Russell's typing of objects is apt to seem somewhat mysterious (unless perhaps viewed constructively as based on a pregiven domain of entities, a position uncongenial to Russell's realism). The source of the problem seems to be that in order to restrict the scope of his logic (consistently with the view that he is still doing logic) Russell needs to delineate the limits of sense or meaningfulness. But he can't do this merely by appealing to the logical nature of the things themselves since this will either assume antecedent grasp of a type of object (most naturally entities) or of a type of object via the applicability of some given predicate. In either case we are trying to motivate limits on meaningfulness from a presupposed antecedent understanding. That understanding is thus exempted from any resulting restrictions of type (just because the restrictions are given in terms of this grasp). Thus either we presuppose adherence to type restrictions (in which case they become inaccessible) or divisions of type threaten to collapse. If, in contrast, the types are shown to correlate with epistemic features we may be able to argue that antecedent grasp of a domain of objects must coincide with a given type.

In *Principia Mathematica* (*PM*) Russell addresses just this question and claims, first, that it is relative typing that is, in practice, of primary importance. That is, in many contexts it is unimportant to determine that actual type of the expression but it is essential to know what type it has relative to another expression. Secondly, Russell resorts to his nominalism to show that the domain of individuals (which, in effect, is simply assumed) does not contravene type restrictions (1910, p.162). We might usefully contrast this logical division based on type with other logical categories such as Frege's distinction between concept and object or Russell's distinction between concepts and things. Now, although these distinction are different they do share an important characteristic in that the "whatever-it-is-in-the-world" wears its ontological status on its sleeve. That is, the ontological categories correspond to the manner in which symbols for the "whatever-it-is-in-the-world" may appear in sentences. So the ontological distinction is only parasitic on *linguistic* competence. Imagine, for instance, what a Fregean might say to someone who maintained that "Frege Russell" was a meaningful sentence.

But these arguments aside, I illustrated in the previous section that admitting a constraint on the generality of logic will be a source of tension with Russell's view of the nature of logic. Thus Russell would anyway have good reason for preferring an alternative solution. A suggestion of a route towards that is to be found in *The Principles* itself.

We arrive at the system of typing by first considering the distinction between singular and plural objects and then extending this simple twofold division by realizing that we can have plural objects of plural objects and so on up the hierarchy. Nothing has been said directly about putative objects which do not fit neatly into this scheme of things, in particular, the nature of propositional functions has been left out of the account. There is a potential worry about propositional functions because the contradiction also occurs in a form relating to propositional functions. That is, if propositional functions can be applied to other propositional functions it seems we can form the propositional function which applies just to those propositional functions which do not apply to themselves. But then this propositional function must both apply and not apply to itself. There is a good reason why Russell does not mention this form of the contradiction in his discussion of the hierarchy of classes and this is that he thinks he has solved this paradox at the end of his discussion of propositional functions,

> It is to be observed that, according to the theory of propositional functions here advocated, the φ in φx is not a separable entity: it lies in the proposition of the form φx, and cannot survive analysis. (1903,p.88)

Although Russell immediately wonders whether perhaps this position, which amounts to denying the termhood of something, leads to contradiction he does tentatively support it. The strategy involves treating the sign for a propositional function as meaningful (in that it occurs in meaningful sentences) although it fails to stand for anything. Thus the statement prefigures the strategy for dealing with the paradoxes which Russell was later to adopt. For what we have here is in essentials the first account of what Russell later calls an incomplete symbol, that is, a symbol occuring in a propositional sign but which does not survive analysis so that the symbol need not stand for a term which is a constituent of the proposition, it has, in this sense, no meaning apart from its use in context.

The attraction of using the notion of incomplete symbols to solve the contradictions is that it promises a means of retaining an unstratified universe of entities. The draw-back is that it requires that we restrict our ontology. But, as we shall see, Russell voices disillusionment with the Meinongian view that every object of thought must be an entity and thus has independent reason for searching for some ontological constraints. "On Denoting" articulates ideas that are pivotal in Russell's thought since it repudiates some of the tenets of *The Principles* whilst also providing, in its development of the notion of incomplete symbol, what seems, *prima facie*, to be a means of rescuing other aspects of that programme.

§6 On "On Denoting":

In "On Denoting" Russell levels attacks on both Frege's distinction between sense and reference and Meinong's view that for every expression functioning syntactically as a singular term there must be an object of thought. In doing so he signals a departure from views he held in *The Principles*. The Meinongian view is accused of leading to contradiction since it demands, for instance, that the existent square circle both exists and does not exist. "On Denoting" is thus often read as an attempt by Russell to rescue the meaningfulness of propositions involving possible Meinongian objects (what I mean by this term will be clarified in the ensuing discussion) without having to assume that there are corresponding entities. However simply to read it as such would, as many readers have pointed out (e.g. Hylton (1990), Griffiths (1981), lead to a misconstrual of Russell's primary motives in developing the later theory.

§6.1 Griffiths on Russell's Early Ontology

In discussing the notion of the null set in *The Principles* Russell notes "that a concept may denote although it does not denote anything" (1903, p.73). This being so it would seem clear that even in *The Principles* Russell has no need to subscribe to Meinongian ontological excess. Propositions which are putatively about Meinongian objects can be accounted for in terms of including concepts which fail to denote anything.

However this is not the view which Russell subscribes to at least relative to the null set. If we have a class concept *a* which is false of everything then it would seem that "all *a*'s" is a denoting concept which denotes nothing: what it purports to denote is a class lacking in members, but, in fact, it must denote nothing since a class is to be viewed extensionally, i.e., not only does Russell accept coextensionality as a criterion of identity for classes he takes the class to *be* a numerical conjunction of its elements. Thus any proposition including the denoting concept "all *a*'s" is about nothing. Russell considers interpreting a proposition such as "all *a*'s are *b*'s" as a formal implication, "*x* is an *a* implies *x* is a *b*, for all values of *x*" but rejects this interpretation because he holds that propositions including denoting concepts (formed by *all*, *every* and *any*) are about specific objects or combinations of terms and, although equivalent to the relevant formal implications, are distinct since those are universally valid whilst, to reiterate, the proposition including the denoting concept is about a specific object. The solution that Russell arrives at is to reject the proposition whilst retaining the interpretations which would have been equivalent had the denoting concept actually denoted something,

> We shall say, then, that, of the bundle of normally equivalent interpretations of the logical symbolic formulae, the class of

> interpretations considered in the present chapter, which are dependent upon actual classes, fail where we are concerned with null class-concepts, on the ground that there is no actual null class. (1903, pp.74-5)

Russell's point seems to be that in interpreting a sentence or formula we are entitled, when the class exists to use any of the equivalent interpretations as suits us. However when we have a null class-concept we must give an interpretation which uses a formal implication or gives an intensional interpretation in terms of class-concepts. So no revision of our *symbolic* practice is needed. The complex corresponding, say, to "all *a*'s are *b*'s" contains legitimate propositional constituents but does not express a proposition. (See Griffiths (1981, p.149) for more details of this argument.) The point that Griffiths (1981) goes on to make is that Russell seems finally to have to endorse a Meinongian ontology in order to make sense of propositions asserting negative existentials which cannot be rejected in the manner just described. Griffiths then argues that the theory of descriptions developed in "On Denoting" is irrelevant to the task of eliminating Meinongian objects since that depends on an interpretation of existence. If existence is taken to be a property of properties then no ontological commitment is enjoined whether or not one subscribes to *The Principles* or to the "On Denoting" theory. Conversely, treating existence as a predicate entails that the theory of "On Denoting" is just as ontologically committing as *The Principles* theory. Finally, he claims that the question of how existence should be treated is itself neutral between the two theories.

I am strongly sympathetic to much of Griffiths' argument. Griffiths distinguishes two possible routes to ontological excess. The first he calls the Meinongian route in which it is claimed that if a proposition is about an entity then it must include that entity. The second is the Fregean route which, although accepting that a proposition may be about an object which it does not include (in virtue of including a concept which denotes that object), still results in ontological excess because it insists that denoting concepts cannot be empty. I have given a brief description of Griffith's discussion of the latter. The conclusion was that Russell is forced to adopt a Meinongian ontology not because he travels along the Fregean route but because he is unable to reject negative existential propositions since they are true. The underlying reason for ontological excess is treating existence as a predicate. My worry here is that *prima facie* truth of a proposition containing an empty denoting concept is not taken to be sufficient grounds for not rejecting it: "all chimaeras are animals" is apparently true according to Russell but must never-the-less be rejected. We minimise the violence this is apt to produce in our symbolic system by reinterpreting any formula in terms of, say, a formal implication (x is a chimaera implies x is an animal, for all values of x) which would be equivalent to the proposition were the denoting concept not empty. Now if this process can be generalized to all denoting concepts and if existence is to be treated as just another predicate then the strategy must again be available for negative existential propositions. Griffith's supports his view that negative existentials are crucial by quoting from Russell's brief discussion of being and existence which occurs in the latter's discussion of space (a good way into *The Principles*). Russell is attacking the theory that every proposition concerns something which exists,

> For if this theory were true, it would still be true that existence itself is an entity, and it must be admitted that existence does not exist. Thus the consideration of existence itself leads to non-existential propositions, and so contradicts the theory. The theory seems, in fact, to have arisen from neglect of the distinction between existence and being. Yet this distinction is essential, if we are ever to deny the existence of anything. (1903, p.450)

This passage makes no mention of non-existent entities indicated by denoting phrases. It succeeds in establishing Meinongian objects, in the sense of non-existent, i.e., subsistent, entities, but fails to persuade one that these can be denoted rather than named. All constituents of propositions are entities but, as the argument points out in the case of existence itself, not all entities are existents. Russell admits, in defining a thing to be anything that can be named, to be using names in a broader sense than usual. Things are nameable and things include "particular existents generally" as well as "many terms which do not exist". So terms which do not exist are still nameable. It is thus plausible to read Russell's argument purely in terms of non-existents which can either be *named* or are concepts. This establishes a Meinongian ontology by allowing non-existent entities but retreats from the worst excesses of that theory in not subscribing to contradictory entities, for which we can only form denoting phrases. Indeed I can find no indication that Russell was ever attached to an ontology of that form and, given his unambiguous rejection of certain propositions because they contain empty denoting concepts, it would be hard to attribute such a view to him. But, given his lack of explicitness about definite descriptions in *The Principles*, it is difficult to attribute any specific ontological view to Russell with any great degree of confidence. What this indefiniteness does perhaps show though is that matters ontological were not foremost in Russell's mind. A final consideration against an identification of *The Principles* ontology with that of Meinong is that in "On Denoting" Russell quickly points out that his earlier *Principles* theory is similar in important respects to the Fregean view he is criticizing but nowhere does he make a similar claim about the Meinongian view.

§6.2 A Recommended Reading of "On Denoting":

How then should we read "On Denoting"? The first thing to note is that it is a mistake to read "On Denoting" primarily as an effort to escape an ontological jungle (in this I agree with Griffiths). The focus of attention in the paper is denoting concepts (or phrases or complexes). Russell repudiates both his earlier notion of denoting concepts and his rejection of complexes containing empty denoting phrases. Russell rejects a complex by claiming that it does not express a proposition but he retains various interpretations which would otherwise have been equivalent. Russell may have become disillusioned with this position since it is just not clear what extra a complex requires for it to succeed in expressing a proposition: in many instances it is a purely empirical question whether a complex does or does not express a proposition. The thought here is that once we have the unity of the propositional constituents there is nothing more for the proposition to be. However it is not safe to attribute any precise motives to Russell on the basis only of "On Denoting" and other writings published at around this time. There is, though, a close relation between the theory of descriptions and the method of rejecting denotationless propositions. The connection is this, if Russell was later disillusioned with the notion of rejecting certain propositions he might choose to treat the proposition (or to select and treat one among the propositions) which that method retains, not as a failed interpretation, but as an *analysis* of the original proposition. The cost of this is that we no longer can treat a proposition containing a denoting concept as about a certain object or complex of terms, rather we have a general proposition containing an apparent variable. This, precisely, is the strategy of the theory of descriptions.

So, as a result of the theory of descriptions we admit that there are genuine propositions with empty denoting phrases. However, existence is still being treated as a predicate (see Russell (1905b) where he draws a distinction between existence as a predicate and existence used in symbolic logic as a property of properties). Griffiths claims that this is ontologically committing since the most natural interpretation of "The *a* does not exist" (if we are treating existence simply as a predicate) gives the definite description primary occurence, i.e., "There is one and only one *x* such that *x* is *a* and *x* does not exist" (rather than "It is false that there is one and only one *x* such that *x* is *a*"). Griffiths is surely right in pointing out that it is not clear when and why we should analyse the definite description as having primary or secondary occurrence but Russell has made some progress in that he has given a possible interpretation which is not existentially committing.

One consequence of these considerations is that the alternative analysis offered of denoting phrases in "On Denoting" must be appraised relative to how succesful one regards his repudiation of the Fregean/denoting concepts view. I shall not attempt to appraise this argument nor shall I explicate it now. However I shall return to the argument below to consider whether or not it supplies a critique of Russell's theory of descriptions itself. At this point I simply want to note the importance that the argument has for Russell's departure from the Fregean view and that the argument seems to supply a perfectly general attack on the idea that one entity can be about, can denote, another entity or complex of entities.

§7 The Epistemological Role of the Theory of Incomplete Symbols:

In a sense Russell has no theory of meaning. His problem is not with our understanding of language but with our apprehension of complexity, where complexity is an objective feature of the world. A necessary condition, Russell claims (explicitly in the penultimate paragraph of "On Denoting"), for apprehension of a complex entity is acquaintance with its constituents. Russell's interest is thus primarily epistemological. Meaning is a concern for Russell but only in the sense that some propositions may be "about" entities which are not among their constituents. *The Principles* cannot be said truly to offer an explanation of this notion of aboutness, rather aboutness is taken to be a product of the proposition's inculding a specific sort of entity, a denoting concept, which stands in a certain (mysterious) relation to the denoted object. This is the logical relation of meaning. "On Denoting" can be read as an attempt to eliminate meaning. This, at first sight, appears odd since the analysis offered in "On Denotng" is one which allows that a proposition may be about something other than its constituents. So meaning at the level of propositions is ineliminable and stands in need of explanation. Indeed the second paragraph of "On Denoting" explains the importance of this just because it is necessary for us to explain how we can have knowledge about and how we can think about objects with which we have no acquaintance. And, in the penultimate paragraph of that paper, Russell returns to the same issue, content that the theory of descriptions satisfies his epistemological constraint on the apprehension of propositions (or complexes) whilst allowing that propositions may be about that with which we have no acquaintance.

The position which we finally arrive at is one in which we admit a whole range of propositions which we apprehend but which are about entities which do not exist: the existence of the entity concerned may be relevant for determining the truth value of the proposition but is not relevant to our apprehension of the proposition itself. We may coherently be in this position for whole regions of discourse according to Russell. Indeed, as we shall see, we are in precisely this position with regard to mathematical propositions involving classes.

Russell's ontology and thus his logic are *epistemically* constrained. It takes Russell some time to realize the full potential offered by the theory of descriptions (in (1906d), "The Theory of Implication", he makes a closing remark which indicates that he is still hoping for a solution in the form of the hierarchy sketched in *The Principles*) but once he does so the task of giving an epistemological foundation and motivation for the resulting logic becomes acute. The 1913 manuscript appears to have been an unsuccesful attempt to carry this programme through for the logic of *PM*.

I want now to look in more detail at Russell's atttempted solution of the paradoxes in terms of his "No Classes" theory and to show how that solution draws support from epistemological considerations. When Russell first moots the theory in (1906a) he notes,

> ... the theory is constituted merely by abstinence from a doubtful assumption, and thus whatever of mathematics it permits us to obtain is indubitable in a way which anything involving classes or relations cannot be. (1906a)

Russell seems to think that we cannot know whether or not classes exist, we can only know certain (non-existential) propositions about classes. This view about our lack of epistemic contact with classes is a consistent aspect of Russell's philosophy. In *The Principles* we precisely need denoting concepts to take the place of classes in propositions because we cannot be presumed to have acquaintance with classes, although we can have acquaintance with propositions which are true of or are about classes. Later, however, Russell makes this epistemological doubt a motivating factor in his treatment of class terms as incomplete. Cartwright (1987, p. 116) notes that the theory of incomplete symbols results in a much "thinner" notion of "aboutness". In *The Principles* something may be the subject of a proposition either if it occurs as a term in the proposition *or* if it is denoted by a denoting concept occuring in the subject position of the proposition. Cartwright claims that an upshot of the theory of incomplete symbols is that, in an important sense, a proposition for Russell can no longer be about anything other than a constituent. The task of the theory is to limit our mathematics to propositions which do not commit us to holding that classes actually exist. The proposition may still be about, e.g., the denotation of a definite description, but only in the sense that that entity is a constituent of a true proposition if the original proposition is itself to be true. Cartwright notes that this consequence should not be rued since the original notion of aboutness was not appropriately linked to intentional properties of the proposition, i.e.,it "implied nothing about the beliefs of one who entertained the proposition" (1987, p.116). The point I want to make here is that as a consequence of "On Denoting" Russell's view of the nature of propositions and hence his logical theory is governed much more strictly by those relations of acquaintance which Russell thinks we can legitimately ascribe to ourselves.

The technical development of the "No Classes" theory depends on the notion of substitution. Rather than using class symbols or symbols for propositional functions as separable entities (an assumption which Russell, at this stage, claims is tantamount to the assumption of classes) we must consider substitutions of one *entity* for another in a proposition. (Note that the idea is *not* that one linguistic item is substituted for another, the notion of substitution is objective.) Thus we have the symbols $p(x/a)!q$ and $p/a;x!q$ both of which mean that q results from p by substituting x for a. (e.g., if p is $f(a)$ then q is $f(x)$.) We can then write q simply as $p(x/a)$ or $p/a;x$, i.e., as the result of substituting x for a in p. So by concentrating on a proposition and the replacement of one of its constituents we can think of the propositions that would result by appropriate substitutions. Thus p/a is called a matrix and plays a role analogous to a propositional function in that it is used to determine a class. Note the the symbol p/a is incomplete since it simply means "the result of substituting a in p by" and this needs completion to make sense. We can, however, use the expression to define a relation of membership. Thus "x is a member of p/a" is defined to mean "the result of substituting a in p by x is true" (See Russell (1906b, pp168–71)), i.e.,

$x \varepsilon p/a \equiv_{df} p(x/a)$ is true.

Having given a contextual definition of the relation of membership we can then simply enough stipulate that matrices should be equal just when they have the same members, i.e.,

$p/a = q/b \equiv_{df} (\forall x)(p(x/a) \equiv q(x/b))$.

This definition has a natural extension to matrices of more than one substitution place. Classes of entities (i.e., individuals and propositions) can thus be treated as matrices with one substitution instance which themselves only have a meaning in use. Thus class terms are not supposed to stand for subsisting entities. Dual relations in extension can be treated as matrices with two substitution places, and so on for relations of higher order.

Classes of classes can also be treated in terms of matrices. Russell gives the following account,

> "The matrix $q/(p,a)$ is called a *class of classes* if, for all values of r, c, r', c', provided $r/c = r'/c'$, then $q/(p,a)\text{;}(r,c)$ is equivalent to $q/(p,a)\text{;}(r',c')$." (1905b, p.176)

and,

> the class p/a is a *member* of the matrix $q/(p_0,a_0)$ if not only $q/(p_0, a_0)\text{;}(p,a)$ is true, but also, whenever $p'/a' = p/a$ $q/(p_0,a_0)\text{;}(p',a')$ is true. (1905b, p.178)

So a matrix, p/a, is a member of another matrix, $q/(p_0,a_0)$, just when the matrix which results from the appropriate substitutions is true *and* when all matrices coextensional with p/a similarly result in true propositions when substituted in $q/(p_0,a_0)$. So, for example, we might have q as "There is at least one nightingale in Berkley Square"; p_0 as

"Pope John Paul is a nightingale in Berkley Square"; and a_0 is Pope John Paul. Then the class *p/a* where *p* is "Plato is an egg in the basket" and *a* is Plato will, provided there is indeed an egg in the basket, be a member of the class of classes $q/(p_0,a_0)$ (just because in this case "There is at least one egg in the basket" will be true.)

Note that the definition of identity of matrices only makes *sense* when the matrices concerned have precisely the same number of substitution places. Thus the defintion of a class of individuals only makes sense when applied to matrices with *one* substitution place and, further, the class of classes must have *two* substitution places. Similarly a class of classes of classes must have *three* substitution places, and so on up (what can now be seen as) the hierarchy. The effect of this typing is to rule out the class paradoxes; it is simply senseless to consider, say, a class as being a member of itself since it can only be a candidate for membership of a matrix containing precisely one more substitution place than it itself has.

Importantly this typing of matrices does not need to be explicitly stated since it is simply a consequence of treating matrices as imcomplete symbols, i.e., of the contextual definition of classes. Russsell, after having outlined the theory, concludes,

> Thus where matrices occur, significance demands homogeneity of type: this does not need to be stated as a principle, but results in each case from the necessity of getting rid of matrices in order to find out what the proposition really means. (1906b, p.178)

The reason why this is important is made clear by Russell in a slightly later paper, (1906c). This paper differs from (1906b) in two significant and related respects. First, in (1906b) Russell begins by setting out his solution to the paradoxes through the treatment of certain symbols as incomplete. This mode of solution is recommended by Russell's diagnosis of the source of the paradoxes as resulting from making certain "false abstractions", i.e., treating propositional functions and classes as entities. In (1906c) Russell endorses Poincaré's diagnosis of the source of the paradoxes as resulting from a vicious circle. This is now formulated loosely as a principle,

> "Whatever involves an apparent variable must not be among the possible values of that variable." (1905c, p.198)

Russell now recognizes that a solution to the paradoxes must issue in a theory of the (apparent) variable which meets this constraint. Later Russell argues that a theory which attempts to give a direct, explicit restriction of the range of the variable is bound to fail since any explicit restriction on the variable can be included as a condition in the generalized proposition which must then be seen as containing a proposition of more inclusive range. (e.g., If we had "For all *x*, *Fx*" provided *x* is an *i*, then we could then have "For all *x*, if *x* is an *i* then *Fx*" and in this latter proposition, Russell notes, the variable must have a more inclusive range.) Thus the variable must have an unrestricted range: its limits are the limits of sense. This is reconciled with the Vicious Circle Principle (VCP) by treating certain signs as incomplete. since such signs are not presumed to stand for entities they cannot be presumed to stand for possible values of the variable. Russell recommends the theory in these terms,

> ... to reconcile the unrestricted range of the variable with the vicious circle principle, which might seem impossible at first sight, we have to construct a theory in which every expression which contains an apparent variable (i.e. which contains such words as *all, any, some, the*) is shown to be a mere *façon de parler*, a thing with no more independent reality than belongs to (say) d/dx or $_a\int^b$. For

in that case, if (say) φx is true for *every* value of x, it will be not true but meaningless, if we substitute for x an expression containing an apparent variable. And such expressions include all descriptive phrases (the so-and-so), all classes, all relations in extension, and all *general* propositions, i.e. all propositions of the form "φx is true for all (or some) values of x". (1906c, p.206).[1]

I have quoted this passage at some length because I think it is revealing of Russell's basic motives and because it is relevant to the second feature which distinguishes (1906c) from (1906b), that is the consideration which Russell gives to the *semantic* paradoxes in the later paper. Before looking at Russell's treatment of the semantic paradoxes let me reiterate what I take to be the main aspects of Russell's position here. Russell's acceptance of the VCP requires that his theory issue in a treatment of the (apparent) variable which complies with the VCP. But this compliance must emerge indirectly from the nature of the theory of itself rather than being an explicit imposition on the theory. Russell's means for achieving this in the substitutional account is by treating certain symbols as not standing for entities. In this way the range of the variable is permitted to be unrestricted without thereby enfringing the VCP.

(1906c) makes an attempt to deal with the semantic paradoxes which are passed over without comment in (1906b). What Hylton (1980) calls the

1. It is worth wondering whether Russell has in fact done enough to justify his claim of *meaninglessness* in taking an incomplete expression to denote a possible value of the variable. It just is not clear why it becomes meaningless, as opposed to being unjustified, to take (say) The φ as a value of the variable. The latter would also be more in concert with Russell's suggestion that the theory is motivated only by refraining from making an unjustified assumption (of the existence of classes, propositional functions etc.). A further (related) question is why, if Russell takes this assumption to lead to paradox, he does not take this as a demonstration of the *falsity* of the assumption, i.e., why is he not justified not merely in refraining from believing that (say) classes exist but also in the denial that classes exist.

simple substitutional theory (as opposed to the ramified substitutional theory which we shall meet shortly) takes the universe to consist in adomain of individuals and propositions. Quantification over all propositions is thus deemed a legitimate operation and this can quickly be seen to lead to development of the semantic paradoxes, For instance we could have,

$(\forall p)(\varphi p \rightarrow p$ is false$)$ $\quad (A)$

Once we have fixed φ A is a proposition which thus falls within the domain of its own quantifier. We can pick φ such that it is uniquely satisfied by A itself. We then have a paradox: assume that A is true. Then, in particular we must have $\varphi A \rightarrow A$ is false. We know that φA (by the choice of φ) so we must have A is false. Conversely, assume that A is false. Now we know that $\varphi p \rightarrow p$ is false is true for all values of p different from A because, by the choice of φ, φp is false for all values of p other than A. So if A is to be false we must have not-($\varphi A \rightarrow A$ is false). i.e., we must have not-(A is false), i.e., A is true. So A is true iff A is false. Nothing in the simple substitutional account prevents the construction of A (and variants).

In the ramified theory Russell distinguishes between elementary propositions which are *bona fide* entities and generalized propositions (or statements) which are not entities. The latter are to be treated only as *façons de parler* and thus cannot be taken to fall within the scope of their own quantifiers,

> Such statements as "Whatever x may be, $x=x$", or "For all values of x, $x=x$", I take to be an ambiguous statement of any of the various propositions of the form "$x=x$". There is thus not a new proposition but an unlimited undetermined choice among a number of propositions. (1906c, p.207)

Thus Russell wants to treat expressions for generalized propositions, on the model of incomplete expressions, as failing to stand for anything. But his means for achieving this similar end differs in either case. Exploitation of the technical device of incomplete expressions requires that there be a suitable context of use in which the meaning of the expression can be explained. Russell, at this stage of his career, sees no context in which to explain the meaning of generalized propositions (which after all function similarly to elementary propositions in terms of being capable of truth and falsity and in being used in inference: both forms of expression seem to be linguistic units possessing their own integrity). Russell takes generalized propositions to be " ambiguous statments" which determine an unlimited undetermined choice among a number of propositions". Later (1906c, p.208) he talks of a generalized proposition as "ambiguously denoting" its instances. But these pronouncements are mysterious. Surely we are not supposed to view such psychological sounding phrases as "choice" *as being* psychological. That is, I do not think Russell can be taken to be suggesting that the generalized proposition is simply an instruction or licence to make some choice. What we are entitled to expect from Russell is an account of the *logical* relation holding between the generalized proposition (whatever it is) and its values. To take the psychological route would be either to ignore the logical relation or to make that relation depend on a psychological relation. Russell would then have to relinquish his belief in the objectivity of logic.

Russell never gives an adequate account of this relation of ambiguous denotation. We can get an idea of why this may have been so by asking what means are at Russell's disposal to explain the relation of ambiguous denotation? Two methods suggest themselves, both, however, would have been anathema to Russell at this time. The first is to see ambiguous denotation as an instance of a certain sort of denoting concept. But clearly, having already repudiated the theory of denoting concepts[1],

1.In the next section, though, I shall be questioning the extent of this repudiation.

Russell was in no position to make such a move. The second strategy is suggested by the later account of propositional functions in *PM* where Russell states that an assertion of a propositional ambiguously denotes its values (*PM*, p.17). That is, the tactic would be to make the account of the apparent variable dependent on that of the real variable. But this would be to resurrect propositional functions. (Perhaps this is part of the explanation of Russell's treatment of propositional functions in *PM* as real.) However there is no way that Russell in his substitutional phase could accept such a solution.

Hylton (1980) points out that the treatment of (all, not simply generalized) propositions as incomplete (as only having a meaning in the context of judgement, belief etc.) cannot have been the motive for Russell's abandonment of the ramified substitutional theory since even in his (1908) theory of types Russell still treated propositions as complexes. Rather, Hylton claims that the reason for Russell abandoning the theory results from more internal tensions between, on the one hand, having to "acknowledge distinctions of type among the propositions which the theory assumes to exist" (1980, p.26), and Russell's argument for a single, unstratified universe which forms the range of the variable. Hylton's criticism is, in one sense, moot since Russell does have to accept that we can legitimately quantify over statements of a given type (where the type is determined by the number of apparent variables in the statement) thus forming a statement of a higher type. Witness this remark made in the course of explaining away the liar paradox,

> [A man who says "I am lying"] cannot mean: "I am now making a statement which is false", because there is no way of speaking of statements *in general*: we can speak of statements of propositions, or statements containing one, two, three ... apparent variables, but

> not of statements in general. If we want to say what is equivalent to "I am making a false statement containing n apparent variables", we must say something like: "There is a propositional function $\varphi(x_1, x_2, \ldots x_n)$ such that I assert that $\varphi(x_1, x_2, \ldots x_n)$ is true for any values of $x_1, x_2, \ldots x_n$, and this is in fact false". This statement contains n+1 apparent variables, namely $x_1, x_2, \ldots x_n$ and φ. Hence it does not apply to itself. (1906c,p.208)

So the ramified substitutional theory does require that we have variables which range over propositional functions or, better, since these are explained in terms of propositions and statements, which range over statements. So Hylton is partly correct since the practice of the theory seems to involve us in recognizing such expressions as expressions for entities. However it is also clear that Russell explicitly denies that statements are entities. His problem is though that he cannot explain their use without treating them as entities. Here the problem about the lack of an explanation of ambiguous denotation recurs.

So, although it seems that this picture presents us with a single domain comprising of individuals (simple entities) and propositions, the reality of the situation for the purposes of quantification is that we have a stratified domain. At each stage we may legitimately quantify over statements which include a specified number of apparent variables. But in doing so we form a statement of a degree one higher than that over which it quantifies. Thus we avoid the paradox but we also develop the need for at least a countable number of distinct (apparent) variables.

§8 The Variable as a Denoting Concept:

The theory of types (1908) which is the next technical development of Russell's programme attempts to limit systematically the range of variables. This should be seen as a distinct second strand to the programme. On the one hand we have the treatment of certain symbols as incomplete and, on the other, we have the typing of variables, i.e., the restriction of variables implicitly to a range of significance. The investigation of the substitutional theory showed that the first feature will not suffice in any simple way to motivate the second. Instead *each has to be seen as arising out of the same epistemological restrictions on the scope of logic.* The programme ultimately fails because Russell's attempt to motivate type restrictions from an epistemological point of view in the 1913 manuscript uses the multiple relations theory of judgement. Russell intends to use that theory to set limits on what it makes sense to judge: constraints of type would then emerge. But the multiple relations theory itself needs to assume type retrictions in order to arrive at a sufficiently fine grained means of distinguishing between judgements and to ensure that it does not describe nonsensical judgements. I shall return to this below.

It should be clear that the notion of the variable is crucial. This is hardly surprising given that Russell connects logic to generality but is also to be expected in the wake of "On Denoting". The achievement of that paper can be regarded as the elimination of *complex* denoting concepts in favour of the variable. (See Hylton (1990) for discussion of this.) The relevance for Russell's programme of dealing with the paradoxes is that the limit of the strategy of treating symbols as incomplete is reached with the variable (which is thus in some sense the ultimate denoting concept). This calls for the second strand of Russell's programme: the range of the variable must be limited to the range of significance of expressions in which it occurs. But, conversely, it seems that the nature of the variable must itself set limits on our understanding of complex expressions or on our apprehension of complexes. The point is that on the one hand we want to say that the range of the variable is set by the range of significance of propositional functions in which it occurs, whilst, on the other we want to say that the type of the variable itself *shows* why certain functions are nonsensical when applied to certain arguments.

There is then an apparent circularity in the programme. It is worth being a little more careful about the nature of the circularity which seems to be beckoning. In a sense, it would serve Russell's purposes were he able to construct a system, which was sufficiently rich to support "enough" mathematics, and which through built in, internal restrictions on the variable prohibited the contradictions. The circularity just alluded to would then only affect the informal motivation for the theory. The circularity could then be interpreted as relating to an ultimate ineffability or primitive of logic. Russell's practice, even if not explicitly informed by such a view, does seem to cohere with it since he never argues but simply stipulates that the type restrictions set up in the logical system coincide with ranges of significance. Russell's writings at this time are dominated by technical development of the theory of types combined with sketched attempts to motivate the theory and promissory notes to carry out that philosophical enterprise. Russell concludes (1908) by noting that the development of the theory of types leaves open a number of philosophical questions about its interpretation. He anticipates a future examination of those questions without being more specific about the questions themselves. So we have some reason to suppose that Russell was still very concerned about giving an informal motivation for his system which is capable of standing up to careful philosophical scrutiny.

I have claimed that the variable is a denoting concept. In "On Denoting" Russell argues quite generally, it seems, against his earlier view and the Fregean view which require a distinction between meaning and denotation. (Here the meaning of a denoting phrase is the denoting concept, the object denoted is the denotation. So think of meaning as (similar to) Fregean sense and denotation as (similar to) Fregean reference.) I must now show how the variable itself does not fall foul of this argument.

The nub of Russell's argument is that we cannot speak about the meaning of a denoting phrase since in any attempt to do so we face a dilemma; either we collapse the meaning/denotation distinction or we loose sight of the *logical* connection between meaning and denotation which thus becomes wholly mysterious. Take a denoting complex, *C*, if *C* is a constituent of a proposition then the proposition is not about the meaning of *C* (or, equivalently, is not about *C* - since what we are considering are meanings which may or may not have a denotation). Such a proposition is about the denotation of *C* (if any). To be able to speak about the meaning of "*C*" we need a new expression (such as "the meaning of '*C*'" or "*C*") which denotes the meaning of *C*. But now *C* cannot be a constituent of this expression (a fact which we recognize informally by using quotes) since then the proposition is not about *C* but about its denotation and there is no way of working back from denotations to meanings since any entity or object may be denoted by indefinitely many denoting phrases. Also we cannot take it that forming a denotation for the meaning by using a device such as quotation marks unproblematically allows us to illustrate the relation of denotation since first, this relation presupposes the relation of denotation which we are trying to explain. And, secondly, it only seems unproblematic because of the linguistic relation between the two denoting phrases. But this is to make the connection "purely linguistic through the phrase", that is, it ignores the *logical* relation of denotation. (See Blackburn and Code (1978) for a good reconstruction of the argument.)

I do not intend to appraise this argument for two reasons. First, the validity of the argument is not my interest: I am interested in whether or not it provides a criticism of Russell's use of the variable. Secondly, some Fregeans (i.e., adherents of some version of the sense/reference distinction) have accepted the argument as showing that it would be impossible to give a direct specification of sense. (Dummett (1973, p.227) and Evans (1982) exemplify this sort of Fregean.) Instead Dummett (with Evans's approval) makes use of Wittgenstein's distinction between saying and showing in order to give content to the notion of sense: the manner in which we specify the reference of the expression within the semantic theory shows its sense. Regardless of whether or not this is an adequate response to Russell's doubts it is clearly one which Russell at this point did not even begin to consider.

For Russell's argument to go through we need to have, first, a distinction between meaning and denotation, (i.e., we need to have one entity which stands in a given relation to others such that any proposition including that entity is about those others) secondly, we must accept that there is no route back from denotation to meaning. Most denoting concepts precisely satisfy these two conditions. The variable, however, does not.

A proposition containing the variable is indeed not about the variable itself but says something about the possible range of values of the variable. So we say (or Russell says) that the variable (ambiguously) denotes any of its values. But the identity of the variable is fixed by its range of values, that is, variables sharing the same range of values are identical. Thus in the case of the variable we do have a route back from denotation to meaning. Russell's argument depends essentially on the complexity of the denoting concept since if we admit complexity we must admit that any object can be denoted in indefinitely many ways (so that there is no route back from denotation to meaning).

§8.1 Hylton's Interpretation and Complex Denoting Concepts;

Hylton (1990) draws attention to the importance of the theory of descriptions in eliminating non-propositional complexity (so that the variable becomes fundamental) but claims that "[t]he arguments against the theory of denoting concepts ... apply equally to simple (non-complex) denoting concepts, such as the variable" (1990, p. 263). In this I claim Hylton is mistaken. Hylton persistently argues that there is a deep tension between what he terms Russell's object based metaphysics (which requires that each entity is what it is independently of all others: that all relations are, in the idealist's sense, external) and a three stage analysis of language in which we have a sentence which expresses a proposition but which is not about this proposition but about some further aspect of reality. Hylton presumes that Russell is dissatisfied with the notion of denoting concepts because it requires that one entity stands in the denoting relation to others. But it is hard to see the source of Russell's difficulties in such simple terms because this is such an overt aspect of what denoting concepts are that Russell surely would not have mooted the idea if it was in such deep tension with his underlying metaphysics. Secondly the aboutness relation is ineliminable at the level of propositions and will thus have to be accounted for in terms of a denoting relation of some propositional constituents. Lastly, the argument Russell gives against denoting concepts makes no mention of any metaphysical suspicions he may have about denotation. True, it does often seem that Russell sees his argument as demolishing the very distinction between meaning and denotation. He concludes the argument, "the whole distinction of meaning and denotation has been wrongly conceived" (1905, p.50). But the argument is, in fact, framed entirely in terms of denoting *phrases* and *complexes* and thus would seem, at least implicitly, precisely to exempt the variable. The problem is that at this point Russell had decided to treat the variable as fundamental but was also unsure of how to account for it. Thus it would not have occurred to him that he should also be considering a meaning/denotation distinction relative to the variable. The argument is simply directed at denoting complexes and makes assumptions specific to denoting complexes.

§9 A Tension in Russell's Attempt to Motivate his Later Theory of Types:

In (1908) Russell defines a *type* as the range of significance of a propositional function. Russell arrives at this definition by first endorsing the idea that contravention of the VCP is responsible for the fallacy involved in the paradoxes. The net effect of this is that certain putative totalities are regarded as illegitimate, i.e., no apparent variable can be taken to range over them. Secondly, Russell goes on to examine the nature of real and apparent variables and of the meaning of generalized propositions. The outcome of this investigation is that we cannot credit ourselves with grasp of a single unrestricted variable but neither can we give an explicit restriction of the range of the variable (since any such restriction can be incorporated within a conditional propositions which can then be seen -if the restricition had any force at all- to have a more inclusive range). (See §7 for more elucidation of this argument.)

In *PM* Russell is less explicit about linking a type to the range of significance of a propositional function. A possible reason for this is that in (1908) there is an evident tension in defining the type in terms of the range of significance of a propositional function. This tension arises because Russell takes the restriction to affect the range of apparent variables. Thus we can no longer enunciate, say, the law of excluded middle as $(p)(p$ is true or p is false) since this involves quantification over an illegitimate totality. Russell attempts to alleviate this difficulty by exploiting the distinction between *all* and *any*, i.e., the distinction between apparent and real variables. If we assert any value of "p is true or p is false" we assert not a new proposition but any value of a propositional function since in this statement p is a real variable (1908, p.67). But, if we quantify over types and if types are limited by the range of significant arguments of a propositional function then we *should* be able to quantify over the range of significance of "p is true or p is false", i.e., we should be able to quantify over all propositions.

In (1908) Russell provides an alternative route out of this difficulty. It is only this latter means that is exploited in *PM*. Here Russell uses the notion of systematic ambiguity of the notions of truth, falsehood, "or", "not", etc. Russell says that it may appear that these may be held to apply to any proposition of whatever order (the order of a proposition is determined by the types of apparent variables which are included in it, presently, I shall outline the hierarchy of types) but this impression is false. These notions need to be disambiguated relative to propositions of a given order. Now, however, it appears that types cannot be defined in terms of the range of significance of a propositional function precisely because we require the notion of type to disambiguate a class of (vitally important) propositional functions. Of course, a type will coincide with the range of significance of a propositional function but the latter notion cannot be used to motivate or define the former. Our present concern is with how the hierarchy of types is established.

§9.1 Individuals and Universals:

Individuals are the fundamental type within the hierarchy. Russell introduces the notion of a *matrix* as a function (of any number of variables) which involves no apparent variables. Our first type consists of all those matrices whose variables range over individuals (eg. $\varphi!(x)$, $\psi!(x,y)$, etc.) plus all those functions formed by generalizing on one or more (but not all) of the variables (eg. $(y)\varphi(x,y)$, $(\exists z)(y)\psi(x,y,z)$ etc.). The next level of the hierarchy includes, first, all those matrices which include variables for first order functions (and, perhaps, individuals) (eg. $F!(\varphi!\hat{x},y)$), secondly, it includes functions derived from such matrices by generalizing on one or more (but not all) of the variables (eg. $(y)F(\varphi!\hat{x},y)$, $(\exists\varphi)(z)G(x,z,\varphi!\hat{y})$ etc.). Note that second order functions may be either functions of first order functions or functions of individuals (or both).

In *PM* Russell gives two informal presentations of the hierarchy of types. In the second presentation (which I have just sketched) a predicative function is identified with a matrix, but in the first presentation a predicative expression is taken to be a function whose order is one higher than that of the highest order of its arguments, so that $(\exists y)F!(\varphi!\hat{x},y)$ is predicative in this latter sense, whereas $(\varphi)F(\varphi!\hat{x},y)$ is not.

I shall be concentrating on two aspects of Russell's theory of types. First, I want to look at the fundamental ontology involved in the theory of types, in particular, I want to consider the constitution and specification of the basic domain of individuals. Secondly I shall want to look at the role of the VCP in motivating the theory of types.

It is unclear what Russell took to be included in the domain of individuals. Individuals are derived from elementary propositions, i.e., from propositions including no apparent variables. So our grasp of individuals is only as precise as is our grasp of elementary propositions. Individuals are introduced as non-functional constituents of elementary propositions,

> The terms of elementary propositions we will call individuals. (1908, p.76)

> ...the terms of ⌊elementary propositions⌋, other than functions, we will call *individuals*. (*PM*, p.161)

The vocabulary used here is reminiscent of that used in *The Principles*, indeed in (1908) Russell refers to the relevant section of that work in a footnote. Recall that in *The Principles* a term of a proposition is taken as the logical subject of the proposition, a thing is said always to occur as a term in a proposition, whilst a concept is seen as having a dual nature in that it can both occur as the logical subject of the proposition and be asserted of the logical subject. Thus according to the above definitions it would seem that a thing or particular is an individual whilst a concept or universal either is or is not an individual according to whether or not it is identified with a propositional function. There seem to be equally serious problems attending either option.

Assume, then, that universals are propositional functions. Now since,

> The universe consists of objects having various qualities and standing in various relations. (*PM*, p.43)

functional expressions must stand for subsisting entities (qualities and relations). So such symbols are no longer to be regarded as incomplete and functions are genuine constituents of propositions. An analysis of the proposition, say, "Socrates is mortal" will show that the proposition contains Socrates and mortality. But the latter is now being identified with the function "$\hat{x}$ is mortal" so apprehending the proposition must involve (if the Principle of Acquaintance is granted) grasping the complex as a value of the propositional function. This is in marked tension with Russell's account of the VCP.

To see why, consider the relation of a propositional function to its values. Russell notes that since a propositional function ambiguously denotes any of its values no value of the propositional function must presuppose the function itself. "This is a particular case but perhaps the most fundamental case of the vicious circle principle" (*PM*, 39). But on the above account grasp of the proposition seems to rely on grasping the proposition as a value of a certain propositional function. So the definiteness of the propositional function requires the prior definiteness of its values which in turn require the prior definiteness of the function. The circularity is evident.

It might be thought that this view oversimplifies the position because it assumes that there is only one "mode" in which an entity may be a constituent of a proposition. One might want to counter the threat of circularity by holding that the function is not a constituent of the proposition in the same sense as is a term, so that grasp of the proposition does not require acquaintance with the function in the same way as it does of the term of the proposition. Indeed in the 1925 introduction to the second edition of *PM* Russell seems to have moved in this direction by exploiting Wittgenstein's distinction between components of a proposition and constituents of a proposition: the relating relation of the proposition is a component of the proposition, the terms of the relation are constituents of the proposition. Whatever use this notion may be put to it clearly is not a distinction Russell is able to exploit to any effect whilst he is wedded to his Multiple Relations Theory of Judgement.

The Multiple Relations Theory of Judgement is an attempt to treat propositional expressions as incomplete. Instead of treating propositions as objective complexes it gives an account of the use of the propositional sign in the context of assertion, judgement, belief, etc. without supposing that the proposition itself exists. Rather than give an account of these propositional attitudes in terms of a two place relation holding between a judger (believer etc.) and a proposition we now are given a multiple relation holding between the subject and the constituents of the proposition he judges (believes etc.). So, for instance, "A is similar to B" is analysed as,

$$U\{S, A, B, \text{similarity}, R(x,y)\},$$

where $R(x,y)$ stands for the (logical) form of a two place relation "Something stands to something in some relation". Prima facie it would seem from this that Russell *did* distinguish between functions and universals. Why else would he use the expression "similarity" rather than simply give the function? Also if Russell thought that the function was identical with the universal it is not clear why he did not simply take it that the function itself determined the logical form of the proposition. Whatever the correct answer is to those questions it is evident that *if* Russell identified functions with universals then the Multiple Relations Theory is inconsistent with the VCP. The Multiple Relations Theory supposes that we have acquaintance with the elements of the analysed proposition. This is why, for instance, Russell wants to broaden his notion of acquaintance to include self-evident acquaintance with logical forms. We make a judgement of a proposition by subsuming the elements of the proposition under the relation of judging. But this must involve grasping the proposition as a value of the propositional function. This, together with the VCP, resurrects the circularity I described above.

The problem here is that in claiming that a propositional function is only definite once its values are definite we are implicitly taking it that those values, i.e., those propositions, are complete in themselves (that they are objectively independent). But this is in sharp contrast with the Multiple Relations Theory. We can only resolve this tension if a propositional function is not taken to be a constituent of the proposition. For then, we can allow that a propositional function "presupposes" its values *and* give an analysis of the proposition in terms of universals, particulars and logical form. But this depends on an ultimate distinction between functions and universals.

In view of the fact that individuals are defined to be non-functional terms of propositions it now seems that universals are a certain sort of individual. Wittgenstein, in his *Notebooks on Logic*, criticizes the system of *PM* as thus interpreted. For, on this interpretation, the theory of types allows us to assert nonsense,

> ...if I analyse the proposition Socrates is mortal into Socrates, mortality and $(\exists x,y)\varepsilon_1(x,y)$ I want a theory of types to tell me that "mortality is Socrates" is nonsensical, because if I treat mortality as a proper name ... there is nothing to prevent me to make the substitution the wrong way round. (Wittgenstein, 1961, p.122)

In the same passage Wittgenstein offers an alternative analysis of the universal as a copula (which is then taken as being simple). He objects to Russell's theory of judgement because it requires one to distinguish between different Types of things - so that for Russell although mortality and Socrates are fundamentally different (in terms of their logical grammar) both are denoted by proper names. The consequence of this within the theory of types is that it allows one to judge nonsense. Thus as far as Wittgenstein is concerned it fails even the most basic adequacy condition on a theory of judgement. (See Griffin 1980)

There are further problems with the theory of judgement itself. These problems were grave enough to have brought Russell's work on the 1913 manuscript to a halt. However, a sketchy version of the theory survives into Russell's Logical Atomistic period (see 1918, pp.224-6). He retains the belief that when one judges or believes something false the relation of judgement or belief does not relate the judger or believer to a complex object but that the relation holds between the subject and the constituents of his belief or judgement. The major problem with the theory concerns Russell's notion of logical form. For Russell the logical form of a proposition is a generalized existential proposition. So, in inferring from a proposition of one form as premiss to a proposition of another form, e.g., *aRb*v-*aRb* from *aRb*, we need to introduce as an extra premiss the form of the inferred proposition. (See Pears (1989).)

I should, for the sake of completeness, consider one final view of Russell's treatment of universals. Linsky (1988) attacks Cocchiarella's (1980) view that Russell identifies functions with universals. The reasons Linsky gives for his view are, in the main, questionable since they relate to Russell's writings well after the 1913 manuscript. But we need not mind this since we have already found reason enough to be dissatisfied with that position. Linsky, however, also rejects the opposing account (championed by Griffin (1980)) according to which universals are individuals. The position favoured by Linsky accepts that universals are fundamental to the ontology of *PM* but maintains that *PM* provides no logic for universals. This may well be the most coherent interpretation of the system of *PM* (certainly there are serious problems with the alternatives) but it surely lacks credibility in the context of Russell's overall programme.

For Russell logic is global. The only manner of restricting the scope of logic that we have thus far managed to uncover stems from an epistemically driven need to prescind from making certain ontological commitments. Now it may well be a moot point as to whether Russell is succesful in this programme (given my interests this is highly moot), but what is beyond question is that Russell has no way of acknowledging the legitimacy of a certain ontology *and* exempting entities of the ontology from treatment within his logical system. This, for Russell, would be an intolerable position. There would be no way of distinguishing essentially logical investigations from those which presuppose a restricted ontology with which they are especially concerned. Linsky thinks that Russell is in just this position in *PM*. If he is right then *PM* is in acute tension with Russell's basic views of logic. Moreover, it is extremely difficult to see how the system could be ameliorated so as to dissolve this tension.

The above discussion invoked a battery of considerations relating to 1) an attempt to motivate the VCP in terms of propositional functions, 2) the basic ontology of the theory of types and 3) the Multiple Relations Theory of Judgement. It is evident that these aspects of Russell's views rather than forming a mutually integrated whole actually vie against each other. It is tempting to view the problems as stemming ultimately from the Multiple Relation Theory of Judgement (which, as noted, is, for independent reasons, unhappy) so that jettisoning that theory would ease the situation. But the importance of the Multiple Relations Theory is that it shows how to treat propositions as incomplete symbols. Once we embark on that project we raise the question of what terms the proposition is to be analysed into. If we make use of propositional functions then, in conjunction with the VCP, we arrive at a circularity, whilst if we have recourse to universals it is unclear where to fit these into the logical theory. Thus it seems that the mere treatment of expressions for propositions as incomplete symbols is the source of the difficulty.

§9.2 Truth and Falsity:

Russell, however, cannot but treat propositions as incomplete since he uses the incompleteness of propositions to ensure that *propositions* do not fall within the range of the individual variable (and thus the semantic paradoxes are avoided). Thus the incompleteness of propositions is crucial in Russell's account of the basic type within the hierarchy of types. Witness this remark,

> We may explain the individual as something which exists on its own account; it is thus obviously not a proposition, since propositions ... are incomplete symbols, having no meaning except in use. Hence in applying the process of generalization to individuals we run no risk of incurring reflexive fallacies. (*PM,* p.162)

Compare this with,

> We may define an individual as something which is destitute of complexity; it is thus obviously not a proposition, since propositions are essentially complex. Hence ... (1908, p.76)

Why does Russell make this switch from the complexity of the proposition to its incompleteness? The reason is that by the time he came to write *PM* Russell no longer thought of propositions as complexes. Rather a complex, a fact, corresponds to a proposition just when it is true so, for instance, *aRb* is true just when there is a corresponding complex *a*-in-the-relation-*R*-to-*b*. Truth is now given an explanation in terms of correspondence rather than being treated as a primitive property of propositions. This enables Russell to give an account of false judgements without assuming the subsistence of false propositions. Thus the proposition is not a complex since it is not an objective feature of reality. The incompleteness of propositions provides a means of accounting for false propositions and simultaneously provides a distinction between individuals and propositions.

I have traced Russell's use of the notion of incomplete symbol to trim his ontological commitments. That process seems to have an unproblematic extension to the case of propositions, thus allowing for a treatment of judgement which does not involve the subsistence of false propositions. However extending the use of incomplete symbols in this way *is* a departure from ealier uses of the notion. Originally, as introduced in (1905), the notion is applied to symbols (definite and indefinite descriptions) which, were they to stand for entities, would stand for entities which are of the same "sort" as those for which we have proper names (or those with which we have acquaintance). This use is then extended to cover class terms. The referents of these expressions, if there be any, are of a different "sort" to those entities with which we have acquaintance. The point is that in the first set of cases we are concerned with a putative individual entity with which we lack acquaintance either because it fails to exist or because we have no epistemic contact with it. In the second case we try to avoid the ontological commitments of an entire region of discourse because of possible scepticism about the existence of such entities. We refrain from committing ourselves to the existence of a whole range of entities because we have a general argument to the effect that we lack grounds for supposing that such entities exist. (See Hylton (1990) for more discussion of this.) This marks a shift in the use of the notion of incomplete symbol but is a change that can still plausibly be seen as an extension of the previous use. The primary motivation is still epistemic even if this assumes a sceptical form in the second case.

The treatment of propositions as incomplete is quite a different matter for here we are presented with what purports to be an analysis of the ontological nature of propositions. The crux of the matter is that our account of the notion of truth hinges on the treatment of propositional signs as incomplete. The use of incompleteness relative to propositions is not motivated by an epistemological doubt about our ability to have acquaintance with a certain sort of entity, it is motivated by doubts about the very coherence of supposing there to be such things as false propositions. Or rather (since this perhaps over states the case) our motivation stems from a doubt about the sense in supposing that there are false propositions since that assumption beggars us for an account of the distinction between truth and falsity. We might put the point like this, in previous uses we have simply found a way of refraining from making a doubtful assumption but now we are not simply refraining from assuming that propositions are objective complexes we *are* assuming that they are *not* objective complexes. It is this assumption which delivers the account of truth in terms of correspondence. (The only alternative to this seems to be to give truth itself a reductive account in terms of epistemic properties. Clearly Russell never considered a programme of that sort. But it would anyway be a relinquishment of his basic realism.)

§9.3 The Vicious Circle Principle, Incomplete Symbols and the System of *PM*:

In this section and the following one I shall be looking a little more closely at the system presented in *PM*. My reasons for doing so are primarily diagnostic. That is, I think I have given good grounds for thinking that the system of *PM* is inadequately motivated and want now to look at the manifestation of that inadequacy in the resulting account of mathematics.

Russell gives the following formulation of the VCP

> "Whatever involves all of a collection must not be one of the collection"; or, conversely: "If, provided a certain collection had a total, it would have members only definable in terms of that total, then the said collection has no total." (*PM*, p.37)

One reading of the principle (one particularly ecouraged by the latter version of the statement) is that it rules out impredicative definitions, that is, definitions which define a member of a totality by quantifying over a domain which includes that totality. The underlying intuition seems to be that we could never grasp such a totality since a putative grasp of the totality would, in some sense, neglect impredicatively definable members or, if it ostensibly encompassed such elements would presuppose an antecedent grasp of the domain and thus would be circular. As Russell puts in a footnote to the above formulation in (1908) "When I say a collection has no total, I mean that statements about all its members are nonsense" (p.63). Such totalities are ungraspable, talk of them is nonsensical.

Let us accept that the import of the VCP is that impredicative specification/definition of an entity is illegitimate. It is, however, important to realise that the VCP itself does not immediately constrain the introduction of incomplete symbols to predicative modes of specification. The reason is clear. Such symbols are introduced by *contextual* definitions which are only required to explain the *use* of the symbols in the context of propositions (or judgements) involving them. We prescind precisely from the assumption that the combination of symbols so defined denotes an entity.

So *PM* *14.01 gives the definition (in use) of the definite description $(\iota x)(\varphi x)$ ("the φ") as,

*14.01 $[(\iota x)(\varphi x)].\psi(\iota x)(\varphi x).=:(\exists b):\varphi x.\equiv_x.x=b:\psi b$ Df.

Now the RHS of this definition quantifies over the domain which, provided the expression $(\iota x)(\varphi x)$ has a denotation, includes the entity thus defined. If the definition specified an entity for which we had an alternative definition then it would not be in violation of the VCP simply because the VCP insists that the impredicative definition be our only definitional access. But the definition does not presuppose that we have independent access to the denotation of the definite description since, first, this would render the use of definite descriptions redundant, and, secondly, the existential quantifier may be demonstrably satisfied without our having to produce the satisfying instance. At *20.5 we have the following specification of the condition for treating the definite description as denoting a member of a given class (which, in view of *14.01, is included in the domain over which we quantify in defining the definite description),

*20.5 $\vdash : (\iota x)(\varphi x) \varepsilon \hat{z}(\psi z) . \equiv . \psi\{(\iota x)(\varphi x)\}$

The reason why these considerations fail to show that we have uncovered a violation of the VCP is, as noted, that we refrain from treating the defined term as denoting an entity. Given any proposition including a definite description we can, according to the definition, rewrite the proposition so as to eliminate all putative reference to an entity denoted by the definite description. The eliminability of definite descriptions shows that the VCP is here inapplicable.

Classes are also introduced in *PM* via contextual definition,

*20.01 $f\{\hat{z}(\psi z)\} . = : (\exists \varphi) : \varphi ! x . \equiv_x . \psi x : f\{\varphi ! \hat{z}\}$ Df.

So the application of a function, *f*, to a class term is explained in terms of the application of that function to a predicative function, φ, which is coextensive with the defining function, ψ, of the class. Propositions involving class terms are thus eliminable in favour of existentially generalized propositions making no such putative reference to classes as entities.

Both uses of contextual definition give the definition in terms of an existential proposition which gives a conjunction of two conditions requisite for the satisfaction of the existential claim. The disanalogy between the two cases is that in the descriptive case satisfaction of the first condition, the uniqueness condition, may be allowed to remain a genuine question. In the case of classes the first condition, the existence of a coextensional predicative function, must not be allowed to determine the truth value of the proposition as a whole. If we have no general guarantee that for any function, ψ, there is a coextensive predicative function then it may be that $f\{\hat{z}(\psi z)\}$ is false *and* not-$f\{\hat{z}(\psi z)\}$ is false. In this case $\psi\hat{z}$ would fail to define a class. (Note that in the analogous case for descriptions the description simply fails to have a denotation.) Accepting the axiom of reducibility precisely furnishes such a guarantee. This is given as,

$*12.1 \quad \vdash:(\exists f):\varphi x.\equiv_x.f!x \quad \text{Pp.}$

The axiom has other uses. It enbles us, at least insofar as we are only interested in the extensional aspects of an entity's properties, to replace the attempt to talk about all its properties (which would contravene the theory of types) with talk about its predicative properties without loss of significance. Similarly, the axiom is also needed in Russell's account of identity which he wants to give à la Leibniz as,

$*13.01 \quad x=y.=:(\varphi):\varphi!x.\supset.\varphi!y \quad \text{Df.}$

(Note the need to treat φ as an apparent variable if we are to give a *definition* of identity.) The restriction of the apparent variable to predicative functions which is required by the theory of types does not weaken the axiom provided the axiom of reducibility is in place, since in that case we can easily prove,

$*13.101 \quad \vdash:x=y.\supset.\psi x\supset\psi y.$

The problem in both the above cases is that we want to be able to consider propositional functions defined by means of quantifying over a domain of propositional functions as members of that totality. And this the VCP obviously forbids. The result is that the hierarchy of types for propositional functions has a ramified (as opposed to a simple) structure. This is clear even from the cursory sketch of the theory of types which I gave above. The effect of this ramification is that propositional functions of a given entity occur at all levels of the hierarchy which are higher than the order of the entity itself.

A point which will be returned to later but which needs to be mentioned here is that the applicability of the VCP to propositional functions (and its corresponding inapplicability to cases of contextually defined terms) shows that propositional functions have a very different ontological status. Indeed it would seem that the applicability of the VCP to a given case betokens a *realistic* attitude to those entities. (I use italics because it is often supposed that the VCP can only be justified from some non-realist or constructivist perspective. See Gödel (1944).)

The VCP has its point of application in motivating the ramified hierarchy of propositional functions. The insistence on predicativity of functional expressions renders the development of much mathematics impossible unless supplemented by the axiom of reducibility. Indeed it is questionable as to whether the system allows for the development of an intelligible theory of classes in the absence of this axiom. But although the axiom does not contradict the VCP it could be argued that it is in strong tension with it. For the axiom seems to guarantee the existence of predicative functions access to which is purely impredicative. This is vague. The idea, though, is that we must, in view of the axiom, grasp the domain of predicative functions as including certain functions, of which the only account we are given, is that they are coextensive with functions formed by quantifying over that very domain. But this suggests that in order to grasp fully the nature of the domain we need to grasp functions specified by quantifying over the domain.

Is this not circular?; If not, why not? and; Would this reason provide a general means of undercutting the VCP? These questions cannot be answered until we have an account of what notion of presupposition Russell thinks is harmful and this Russell does not give us. But it is clear that there is no blatant circularity, if only for the reason that we are not supposed to have a grasp of the predicative functions via this route, we only have a means of determining their extensions.

Both Quine and Ramsey recommend rejection of the VCP and thus have no need for the (motiveless) axiom of reducibility. The resulting functional hierarchy is simple. The effect of this is however to require a fundamental distinction between the semantic and set-theoretic paradoxes (a distinction introduced by Ramsey). Russell identified both forms of the paradox as stemming from the same source, or, rather, he did not acknowledge such a distinction at all. This was not simply an oversight on Russell's part. His views on the universality of logic, that anything expressible should be expressible in a single logical system, would have prevented him from adopting the Ramsey/Quine approach.

§9.4 Goldfarb's Account of Ramification in *PM*:

My interest here is anyway primarily diagnostic. So rather than canvass support in advocacy of a particular solution I am questioning the coherence and nature of Russell's basic approach. Warren Goldfarb in a recent paper (1989) on "Russell's Reasons for Ramification" is helpful here. Goldfarb attempts to give some rationale for ramification which trades on the distinction between intensional and extensional items and thus does not import constructivist scruples. (I use the horribly but, in this context, helpfully bland word "item" to avoid the suggestion that we are necessarily discussing entities.) The crucial difference between these two sorts of item seems to be that grasp of an intensional item suffices for grasp of *which* item it is. The identity of an intensional item thus is given by its manner of specification or, better, presentation. Extensional items, in contrast, can be grasped before we grasp *which* item is grasped. That is to say that we may grasp a specification of a given extensional item and still have work to do in determining whether or not it is the same item as that given by another specification. Goldfarb argues that this difference in questions of identity relating to intensional and extensional items enables one to justify ramification for intensional items whilst accepting no ramification for extensional items. Thus the justification offered imports no non-realist metaphysics. The predicativity requirement on intensional items arises from considering the identity of the item as determined (partially) by the identity of the variables it includes. Goldfarb's suggestion is thus that, in his insistence on ramification, Russell is sensitive to these subtle questions of identity conditions and thus betrays no commitment to a constructive view.

In apparent contrast to Goldfarb I have linked the applicability of the VCP (or, equivalently the requirement of predicativity) to the question of whether or not Russell views the symbols concerned (realistically) as denoting entities or as only having a meaning in use. This, however, belies an underlying sympathy between the two views. Goldfarb wants to say that ramification in Russell's system is *not* the product of constructive metaphysics but is consonant with Russell's basic realism. Indeed Goldfarb tries to show that ramification is a consequence of Russell's (extreme) realism with repect to variables which Russell sees as constituent entities of propositions. Goldfarb conjectures also that perhaps Russell saw variables as, in some sense, including their domain of quantification in which case "even the weakest form of the VCP suffices to yield ramification" (1989, p.37). My description also links Russell's ramification to a realistic attitude (although I have offered this only as a piece of exegesis). The difference between the two views thus appears to be that I emphasize the role of contextual definition in countering any tendency towards ramification whereas Goldfarb wants to attribute this to the possession of an extensional criterion of identity.

I think that Goldfarb is precisely correct in bringing to the fore the issue of identity conditions of intensional and extensional items. There is a close relation of this issue to the business of contextual definition and reflection on that, I think, shows that the the two views do come together. An extensional item can be grasped prior to having grasped which item is grapsed. For items introduced by contextual definition we *prove* statements specifying conditions of identity as special cases of propositions involving the relevant incomplete symbols. So for classes we prove, given the axiom of reducibility, that classes are identical just when they are determined by formally equivalent functions (which is a version of coextensionality). Indeed Russell takes the ability of his contextual definition of classes to furnish a proof of this, "the essential property of classes" to be a justification of it. Thus the system of *PM* is given a measure of support through its ability to cohere with antecedent notions of identity for certain sorts of item.

This is not an unattractive position since it shows, at least, that Russell does make a contribution to the study of the logic of extensional items as derived from the laws governing intensional entities. In sum, the observation shows that treatment of certain symbols as incomplete is only a partial motivation for the system of *PM*. We need to use some adequacy constraint on the range of variables and Russell gives this as the VCP. The ramification thus introduced does not however apply to descriptions (since they are incomplete symbols) and cannot be allowed to infect class terms. The axiom of reducibility is introduced to enable a coherent theory of classes, i.e., for the contextual definition of classes to be intelligible and for it to cohere with antecedent grasp of the identity conditions governing classes. Conversely, the ramified hierarchy of propositions and propositional functions is partially justified by the intensional nature of these entities and, in particular, by the need to individuate the variables included in such expressions. So the VCP is introduced as a necessary constraint on the range of (apparent) variables but is tested, insofar as the system as a whole is tested, against conditions of individuation of certain items. This suggests that the direction of argument is from the manner of determining the referent or denotation of an expression (in a suitably benign sense of determining its identity) to a restriction such as the VCP. To know the sense of an expression is to possess information sufficient to identify the referent (in appropriate circumstances). An account of the sense of a category of expressions will thus give a general account of the identity conditions governing the relevant sort of item. The argument thus appears to be from a theory of sense to revision in logical practice. Russell never develops an explicit theory of sense, so his system suffers from lack of principled motivation. Also it is not clear that a theory of sense would endorse the VCP in the form propounded by Russell. (I investigate this question in chapter 5.)

§10 Conclusion:

The conclusion I wish to urge is, first, that the purely epistemic motive for restricting the scope of logic (the need to prescind from cetain doubtful and ontologically committing assumptions) is insufficient. This is because it fails, i) to motivate a principled account of universals; ii) the account of propositions as incomplete symbols is not epistemic but is primarily ontological; and iii) the identity conditions of various sorts of items become crucially important in motivating and appraising our choice of logic. This last point brings me to the second aspect of my conclusion. It is this, the thrust of Russell's work after "On Denoting" is an application in new contexts of notions which are developed in that paper, specifically, of the the notion of incomplete symbols and of the elimination of Fregean sense (the treatment of denoting complexes quantificationally). I argued that the achievement of "On Denoting" admits the variable as fundamental, as, in a sense, the ultimate denoting concept. If this is so then the notion of Fregean sense has not been eliminated. What we have is, rather, an argument in favour of the non-reification of sense and a reductive account of the sense of certain expressions in terms of the sense of other expressions. The preoccupation with conditions of identity over the last few pages is, I think, testimony to the need to consider issues to do with the sense of expressions. Most importantly, there is a lacuna in the programme which must be filled by an account of the sense of variable expressions. Russell never broaches this aspect of the programme (indeed it is hard to see how he could given that he thought all apprehension was based on a primitive relation of acquaintantce with some objective entity) but it is precisely a substantial theory of sense that is needed to give a principled means of restricting the logical practice. In its absence Russell's system, in particular, those twin features of it, namely, the axiom of reduciblity and the VCP, will appear ad hoc. To put the point slightly differently, Russell's "justification" of his account of classes in terms of its ability to prove the extensionality of classes should not be treated merely as extra systematic. The justification may not be formal but it is fundamental to his system and thus calls for an explicit account of sense.

CHAPTER TWO: THE NATURE AND ROLE OF A THEORY OF MEANING

The Nature and Rôle of a Theory of Meaning

§1 The Philosophical Task:

Consider the following provocative description of philosophy. Given a practice our philosophical task is to understand it. When that practice is linguistic we want an account of the meaningfulness of the practice, we want to be told what confers meaning on expressions in the practice. Such an account is desired because it exposes the "workings" of the practice, it shows us the role of expressions in the practice and the role of the practice in language as a whole. In so doing it gives us access to the content of the practice since it results in an explanation of the concepts used.

If our best efforts in this endeavour are frustrated and the manner of our failure suggests that such an account is impossible then we must admit either that we cannot get a clear view of the workings of our own language - that meaning can be, in principle, obscure - or that despite the persistence of a seemingly coherent practice that practice stands in need of revision. The consequent of this conditional is a disjunction of unwelcome conclusions. The first disjunct is unwelcome because it suggests that meanings are indeterminate (or, at least, are only mysteriously determined) and this makes it impossible for us to know what we are saying. The second disjunct incurs an obligation. What more can we take as a criterion of meaningfulness than non-collusive agreement in use? What perspective gives us principles which so enable us to so criticise a practice? That is, we are asked to imagine a situation where non-collusive agreement in use is simply a pre-condition of there being a practice which stands in need of justification. We then want to know what perspective supplies the additional criteria by which we judge a practice to be meaningful or justifiable. For example, although classical mathematics is certainly a practice (since we have no difficulty in deciding whether a construction is a good piece of classical mathematics) it would seem that we can still raise the question of whether it is a justifiable practice.

This sketched panorama begs and raises a number of questions. It begs an important question because it assumes that a complete "external" account of the practice is possible. Might it not be the case that we gain an understanding of the practice only through being inducted into it? Perhaps there is no distinctively philosophical or external way of understanding the practice. It raises questions because we need to know what sort of external account is required; How ambitious need the project be so that success guarantees enlightenment? ; In what terms can we give an explanation?

The business of this chapter is to canvass Michael Dummett's views on the correct answers to these questions. So my aim is limited and is primarily expository. I shall, of course, be considering arguments for the anti-realist position but make no claim to be offering a thorough defence of anti-realism. The task I have set myself in this chapter is to be clear about the main features of the anti-realist view. This coheres with the major concern of the thesis which is not to offer an argument for anti-realism but to investigate the revisionary implications of that position. At two points (in considering McDowell's and Haack's views) I do offer a defence of anti-realism but even there the primary aim is to use those arguments to reflect further on the nature of the anti-realist perspective. I undertake that task because I think both Haack's and McDowell's attacks seem to accept (or pretend to be neutral about) basic elements of Dummett's position whilst criticizing the scope or coherence of anti-realism. So, even at those points, my defence is limited to showing that no good reason has been given for departing from the basis of the anti-realist project (and not, indeed, that we should accept that programme).

I shall turn to the question of revision towards the end of the chapter, specifically, to how much revision Dummett thinks is called for in classical logic and mathematics. For the moment I simply want to note that if a project along these lines is deemed coherent then, unless we have grounds for unbridled optimism about its success, there just is the possibility that the antecedent of the conditional will be fulfilled and so we shall be committed to its consequent. We've no reason, *ab initio*, for imposing as an adequacy condition on an explanation that it succeed in explaining the whole practice. Whether use is or is not sacrosanct will emerge in our examination of the nature of meaning. We have, in advance, no motive for stipulating that meaningful use coincide with established use.

§2 The Form of a Theory of Meaning:

To learn a language is to acquire a skill. A skill for which it makes sense to ascribe acquisition should be one of which we can give an account. To suppose otherwise would be to suppose that constitutive aspects of that skill were ineffable. This is alien to our concept of language use as essentially social since it would become utterly mysterious as to how we reliably ascribe possession of the skill. It is essential that in speaking a language we treat our fellows as understanders of, at least the relevant portion of, language. This presupposes that we can reliably ascribe understanding of any fragment of language.

These thoughts indicate that we have room for an account of what constitutes mastery of a language. How should such an account best be tackled? The project requires an analysis of the concept of meaning. This is done via the construction of a theory of meaning or, rather, this construction becomes a technical device in attaining our ends. The actual construction of the theory is never seriously envisaged but we inquire into what form the theory would take were it to satisfy certain predetermined constraints.

We determine the constraints by considering two related aspects of the explanation: how much explaining need we do?; and, what can we use in our explanation? Our aim is complete perspicuity, i.e., we do not recognize that there may be semantic features which, in principle, evade our account, so we place great demands on the comprehensiveness of the explanation and are sparing in what we use in the explanation. So as a matter of methodology an adequate theory is required to give a complete non-circular account of the meaning of every expression in the language. Dummett gives voice to this as follows,

> What would render the functioning of language unintelligible would be to suppose that the relation of (immediate or remote) dependence of the meaning of one word on that of others might not be asymmetrical, that, in tracing out what is required for the understanding of a given sentence, and, therefore, of the words in it, we should be led in a circle. (1977, p.368)

We want a description of the language which makes it perspicuously learnable so if our theory involves itself in the sort of circularity just described it will be impossible for us to account for how a speaker gains access to the meanings of the expressions concerned. (This position might be modified to take into account certain limited local holistic relations of meaning, e.g., it is plausible to suppose that certain contraries, "child" and "adult", say, might need to be learned together. However a situation is regarded as pernicious if it entails that, in order to grasp a given expression, we need to grasp expressions of *arbitrary* complexity.)

A theorist equipped with such a theory would have complete mastery of the language. The theory is intended as a systematic representation of what a speaker knows when he knows how to speak a language, it is a theory of understanding. We cannot achieve this if the explanation fails to give a full account of what this knowledge consists in or gives-an explanation by recourse to concepts which presuppose the possession of language. In Dummett's terminology the theory must be full-blooded (see Dummett 1975c, p.102). Note, this is not intended as an argument for a full-blooded theory of meaning. The demands just made of a theory of meaning do indeed lead to full-bloodedness but an argument for full-bloodedness would have to take into account what we want a theory of meaning for, it would have to show why a theory which fails to be full-blooded is inadequate - so the only acceptable *theory* of meaning is full-blooded - and that a full-blooded theory of meaning is a possibility. I shall elaborate on the nature and feasibility of a full-blooded theory below.

In contrast a modest theory of meaning (Dummett 1975c, p.102) only attempts a description of our use of language by presupposing concepts expressed by certain expressions taken as primitive. One thought why we might be dissatisfied with this is that these primitive concepts cannot be language specific since, if they were, translation of one language into another (if possible) would enable the primitive concepts of one to be explained in terms of the second and, by reversing the process of translation, vice-versa. This explanatory equivalence shows the concepts to be not genuinely distinct or else that in either language we have a choice as to what concepts we take as primitive so that we could choose to take the same concepts as primitive. Now, if the concepts are essentially linguistic, in the sense that grasp of the concepts must involve competence in using certain expressions, we must gain that grasp via immersion in the practice of using a certain set of primitive expressions. But this, supposedly, can be attained through competence in using the primitives of *any* language. For that to be the case it must be possible to recognize the similar rôle played by the expressions in non-linguistic contexts, i.e., in contexts, say, involving estimation of speakers' motives, purposes or intentions. In other words, if grasp of the basic concepts necessary for language acquisistion is not, implausibly, to be tied to grasp of the use of specific expressions within a given language we must be able to recognize the similarity in the rôle of the primtive expressions in different languages. For this to be possible we must be able to situate this rôle in our general practice of interpreting the behaviour of other speakers, and this exercise is not essentially linguistic. Explanation in the theory of this rôle would however be to give an account of what it is to grasp these basic concepts in terms of capacities which are not essentially linguistic. Thus we would resume a full-blooded account. For example, if nouns for colours were taken as primitive expressions (and the concepts of the individual colours as primitive concepts) in a given theory of meaning then, if we assume that two languages have equivalent expressive power, they must both include expressions for the same colours (even if they differ over whether a given colour is denoted by a complex or simple expression). To recognize that the two languages both succeed in enabling talk about colours would require that we show how in each language colour terms were used by speakers in distinguishing and identifying various appearances of colour. But that account would give an explanation of grasp of colour concepts in terms of our ability to re-identify and discriminate between instances of various colours. One way of doing this is to apply the colour vocabulary. We would thus give an explanation of how a given concept is expressed by a certain expression. Dummett (1978b) considers a similar case in our grasp of the concept "square",

> ...it can only be by reference to [the ability to apply the word "square" to square things and not to others] ... that we can explain what it is to associate the concept *square* with that word. (1978b, p.7)

Note that the theory would no doubt make use of the colour terms of the language it was framed in. This does not frustrate full-bloodedness since we still get an analysis of what it is to exercise a colour concept in terms which do not presuppose that the concept is already possessed.

This deals with the first alternative according to which the basic contents were assumed to be essentially linguistic. So assume, now, that the concepts are not essentially linguistic. This has an acceptable and a controversial interpretation. If it is to mean that the speaker must have a grasp of certain concepts which are prior to and thus independent of a linguistic medium and which provide the basis for attaching meaning to words via some process akin to internal ostension or association then the modest theorist has work to do in explaining why, in view of the fact that we cannot give an account of this, we do not restore a postion which acknowledges an ineffable semantic component, that is, a position where the idea that we share meanings becomes at most a hypothesis. (I shall consider positions of this sort shortly.) But if the thought is that we describe the possession of certain concepts in terms of engaging in the appropriate non-linguistic social practices then we have a full-blooded theory of meaning.

§3 A Theory of Meaning is a Theory of Understanding:

We use language to communicate. What a speaker communicates when he uses a sentence is the meaning of the sentence as uttered by him in that context of utterance. The meaning of a sentence seems to reside just in its rôle in communication and communication is possible only because a sentence's meaning is public, in a sense, open to view. What this means is that the meaning of a sentence resides in its use and has no *essential* component consisting of, say, a mental event accompanying use. (Note that there is no direct argument against the possibility of a logically private language from this position. I am simply assuming, if it can be called an assumption, that we do communicate. The issue about privacy in language is to do with accounting for normative constraints of our use of language in terms of some mental accompaniment of use.) The reason why this mentalistic view is wrong is that it results in it always being at best a hypothesis that we share meanings and this a subversion of the communicative function of language –so a *reductio ad absurdum* of any such notion of meaning: the assumption of mutual understanding does *not* have the character of a hypothesis. If mutual understanding is only a hypothesis then it only becomes reasonable if it can be shown to be an explanation of shared use. But our use is the *sole* reservoir of evidence so, if agreement in use is ever a ground for assuming that we share meanings, then it always is. So that element of understanding which is unavailable in use drops away as redundant. The idea that meanings are distinct from use seems, in view of this, to be contentless. Whilst, if we are not justified in taking agreement in use as agreement about meaning, then we can never have good grounds for supposing that we understand each other. The point here is not verificationist. What is claimed is that if someone believes that meaning does not reside in use and prefers to think of meaning in terms of some mental accompaniment of use then he must show how that accompaniment confers conceptual content on the expression. The problems with this view have been formidably articulated in the early sections of *Philosophical Investigations*. The central problem seems to be that we cannot make sense of the concept as a (recurrent) mental event. Rather all we can think of is an image or picture in the mind of the speaker. And now the supporter of this mentalistic view must explain how the nature of this item can matter in the least in ascriptions of meaning provided that the speaker's use agrees with that of other speakers. Also, the mentalist has work to do in explaining how the image actually connects with use. Although pictures often seem to be self-intimating this is only because we know how they are to be interpreted. But this should not obscure the fact that they do require interpretation; we need to know how they are to be used. Explanation of this would involve us in a regress. Thus the explanatory value of the mentalistic view vanishes.

So accounting for language as a medium of communication demands that we accept that agreement in use be agreement in meaning. For to suppose it is not is to suppose that there is some transcendent fact of meaning about which we may disagree consistent with agreement in all our linguistic behaviour, i.e., that mutual understanding is a hypothesis. So we accept the slogan "meaning is use" and seek some clarification of it.

To give the meaning of an expression is to say what a speaker knows when he understands it. We want an account of what this knowledge consists in. Dummett claims that an ascription of knowledge must, if it is not to fall into vacuity, show how the knowledge is manifestable, i.e., show what distinctive capacity is involved in possession of the knowledge, and how it is acquirable. In support of this view Dummett recalls an example of Wittgenstein's,

> ... a dog can expect his master to come home, but he cannot expect him to come home next week; and the reason is that there is nothing the dog could do to *manifest* an expectation that his master will come home next week. *It makes no sense* [my italics] to attribute to a creature without language a grasp of the concept expressed by the words "next week". (1978b, p.5)

The claim is not simply that we lack a justification for ascribing knowledge (of the meaning of an expression) in cases where the subject is incapable of exercising a capacity which displays that knowledge, but that in such cases we cannot make sense of the putative knowledge ascription. So we have a (metaphysical) constraint on our concept of knowledge.

Manifestation: If a speaker knows the meaning of an expression then there must be no aspect of this knowledge that is not capable of being displayed in the use the speaker makes of the expression. What demonstrable capacities can we take as constitutive of understanding? For many expressions we can be satisfied with an explicit verbal explanation offered or recognized by the speaker, provided that we have reason to believe that the speaker understands the explanation. Clearly, though, on pain of regress or circularity, this cannot amount to a global account of manifesting an understanding. So for some expressions understanding can only be manifested by the use the speaker makes of the expressions in certain circumstances.

The theory of meaning gives a theoretical representation of this practical capacity for correct use. Obviously we cannot ascribe to the speakers an explicit knowledge of this propositional representation for then it would have sufficed as a verbal explanation. The speaker's practical capacity is thus regarded as implicit knowledge of what is expressed in the theoretical proposition. So the proposition can only make use of concepts which can be fully explained in terms of actual capacities of speakers. So, where a sentence describes an effectively recognizable state of affairs, that understanding is manifested in the speaker's ability to put himself in a position to exercise that recognitional capacity and to assent to or dissent from the sentence as appropriate.

Acquisition: Actual, not ideal, speakers gain a mastery of language. When we learn a language we learn from the uses others make of the language and from the explanations they offer of that use. Limited experience of such linguistic use enables the successful learner to go on to make his own novel utterances and to distinguish correct from incorrect use. A theory of meaning must explain how a speaker gains competence using only capacities he can be credited with.

The manifestation and acquisition constraints are often regarded as twin themes in the publicity of meaning (see, for instance, Tennant 1987, p. 4). It is important to be clear about the demands of publicity. We have been given an argument for holding that in order to account for language as a medium of communication meaning must be public. In other words the possibility of communication requires that in our account of language we adhere to the "meaning is use" slogan and we apply that constraint via the manifestation and acquisition challenges. But if this, without further ado, is simply taken to require that all use be *publicly* manifestable it begs the question against the private linguist – obviously the private linguist cannot publicly manifest his grasp of his expression defined by internal ostension; if this is taken as a criticism of the meaningfulness of his use then his task is, *ab initio*, impossible. The notion of a private linguist results from a Cartesianism which seeks to ground language in a private epistemic realm. Our constraints do not immediately rule out that position but require the Cartesian, first, to show the item of private knowledge which he takes to be constitutive or determinative of meaning to have a "hard enough" link with use to ensure satisfaction of our constraints. Secondly, we need to be convinced that a logically private language is so much as possible. Our discussion of language thus far has concentrated on its communicative rôle but we have uncovered constraints on meaningful use *per se*. The point is that manifestation is a general requirement on ascriptions of knowledge and that the private linguist cannot manifest his putative knowledge to himself since any attempt to exercise the skills which constitute possession of that knowledge will count as being successful provided only that the exercise *seems* successful to the private linguist. So he cannot give content to the notion of a mistake in exercising his putative knowledge. Thus he cannot reliably self-ascribe understanding and a private language is not possible. Now we are in a position to motivate a concern with public manifestability as a constraint governing *all* meaningful use.

§4 Full-Bloodedness and Modesty:

The issues of whether only a full-blooded theory of meaning will satisfy our requirements of a theory of meaning and whether such a theory is a possibility are important in their own right but also help clarify the demands of manifestability. McDowell (McDowell, 1987) criticizes Dummett's insistence on full-bloodedness for collapsing into a behaviouristic account of understanding because it requires a complete behavioural description of grasp of a content. McDowell sympathises with Dummett's aversion for a psychologistic account (i.e. an account which relies on giving the meaning of an expression in terms of some relation of the expression to a mental process) because any such account makes the belief in shared meaning a hypothesis. So, if we accept the demands of full-bloodedness the account we give of the grasp of any concept must show how that grasp consists in capacities which are demonstrable in aspects of overt behaviour. Now this postion, McDowell tells us, threatens to leave mind out of the account altogether, i.e., we collapse into behaviourism.

Dummett attempts to use the notion of implicit knowledge as pivotal in giving an account which avoids both the Scylla of psychologism and the Charybdis of behaviourism. Dummett notes that implicit knowledge is "knowledge which shows itself partly by manifestation of the practical ability, and partly by a willingness to acknowledge as correct a formulation of what is known when presented" (Dummett 1978b, p.3). McDowell points out that this notion enables us to avoid psychologism since the manifestation of the practical ability ensures that possession of the knowledge is, appropriately, open to view. So psychologism is avoided, whilst the fact the formulation of the knowledge possessed is recognized as correct shows that the knowledge is functioning in *guiding* behaviour -so we are not simply picking out regularities in behaviour but ascribing minded activity. Thus behaviourism also is avoided.

McDowell has two reasons for claiming that this approach fails. First, he points out that any formulation of what an implicit knowledge of an expression consists in will, in many instances, use the expression itself so that if it is to be recognized as a correct formulation of one's implicit knowledge it will presuppose that one understands the expression correctly. If, however, one objects to this, arguing that the linguistic expression itself is not relevant to the formulation of the implicit knowledge but that what matters is what is expressed, i.e., the *content* of the formulation, then we will have resurrected psychologism -the idea that we can strip away words to get at the pure thought lying behind the conventional means we have for expressing the thought. McDowell's second reservation is that the outward manifestation of the implicit knowledge can never succeed in completely displaying the knowledge possessed by the speaker since the speaker only exhibits finite episodes of use and these are reconcilable with infinitely many formulations of what the implicit knowledge possessed by the speaker actually is.

I think that McDowell is right in focusing on the importance of the notion of implicit knowledge but wrong in his account of how it determines a path between behaviourism and psychologism. First, the idea that we might exhibit implicit knowledge in recognizing a formulation of the knowledge cannot be the whole story of how mind enters our account of understanding. The reason for this is that if we insist that implicit knowledge always be capable of being exhibited in this manner then circularity is bound to infect our account just as it does if we insist that understanding is always explicit knowledge of some verbal paraphrase. Moreover, we often want to ascribe grasp of a concept to, say, prelinguistic children or to animals or, simply, to persons who lack the linguistic resources to recognize a formulation of their putative implicit knowledge. Thus the ability to recognize a formulation of implict knowledge is not constitutive of possessing that knowledge (although, obviously, it will, on occasion, warrant ascription of implicit knowledge). So the point at which our account requires the involvement of mind is not exclusively to be equated with this capacity.

We can, however, salvage this insight from McDowell's remarks. The importance of the ability to recognize a suitable formulation of implicit knowledge was that it linked ascription of implicit knowledge with the idea that a speaker is guided in his practice by this knowledge. It is now obvious, though, that the outward manifestation of linguistic behaviour is important not simply because of its discernable regularities but because of normative constraints which govern the correctness of that behaviour. A full appreciation of this point would show why McDowell's second objection is also misguided since all that that objection reveals is an underdetermination of what the underlying regularity in our linguistic behaviour is, given only a finite sample of that behaviour. We are not interested in regularities of linguistic behaviour but in the correctness of that behaviour.

This point obviously needs a great deal of elaboration. It touches on a theme running deep in much philosophical thought since the early years of the century and permeates much of Wittgenstein's later writings. I cannot ignore it but must keep my remarks on it brief. My interest here is the revisionary consequences of Dummett's programme but, to gauge them, we need to understand the nature of his programme. Understanding a philosophical position cannot be divorced from the sort of argument canvassed in support of it. So what I want to do now is simply to sketch an argument in support of Dummett's notion of full-bloodedness ignoring possible objections to it and refinements that may arise from such objections.

Implicit knowledge possesses an outward manifestable aspect which refutes accusations of psychologism yet the character of that manifestation exhibits the influence of mind so halts the slide into behaviourism. How is this possible? The problem with such a position is that it is perfectly possible to give accounts of behavioural manifestations of understanding which presuppose some conceptual accomplishment but that in doing so one would threaten the requirement of full-bloodedness. (This seemed to be the problem with an account which relied on acknowledgement of formulations of the implicit knowledge possessed.) So what we want is a form of behavioural manifestation which shows the involvement of mind yet which does not presuppose grasp of certain concepts. We need, first, to be clear about what we want of a theory of meaning. "We are not looking for a theory with predictive power, but for a *description* that makes sense of the activity as one carried on by rational beings." (Dummett 1987, p.260) The point that Dummett seems to be making here is that the very nature of our interests in a theory of meaning preclude the possibility of behaviourism since we presuppose we are giving an account of the activity of minded beings. Dummett goes on to amplify briefly what this entails and to describe why it is no concession to modesty,

> An adequate theory of meaning must allow for this process of estimating a speaker's intentions, but should not incorporate a description of it. The process is in no way special to the particular language; it is based upon an understanding of the language, but does not involve anything that has to be learned in learning that language rather than any other. Save for its subject-matter, it involves nothing special to language as such: we estimate the intentions and motives underlying other people's utterances by the same general means as we estimate those underlying their non-linguistic actions. (Dummett 1987, p. 261)

We could, it seems to me, make the same point in the context of translation by claiming that we are never in the position of a radical translator since we bring to every translation situation the same standards of rationality which govern our interpretation of others' behaviour. We impose on the speaker our standards of rationality as the only means of making *sense* of his behaviour. What the indeterminacy of translation shows is that, if we constrain the evidential base for translation to overt uses of language in observable circumstances, we cannot motivate a single scheme of translation as being correct. Our choice of translation manual can then only be made on pragmatic grounds. Were we to possess a complete specification of standards of human rationality then the argument for indeterminacy of translation would be undercut. There *may* indeed still be a measure of indeterminacy in any given translational project (just because our construction of theory is often subject to pragmatic and historical factors) but Quine's *argument* for the conclusion that significant indeterminacy is guaranteed would no longer hold. So complete specification of human standards of rationality would be to give a generally applicable methodology of translation which we have yet to be given reason to suspect is inadequate. (We could view the resulting position as one where Davidson's Principle of Charity, rather than trading on some notion of agreement, is given a completely explicit enunciation.) So accepting the argument for indeterminacy of translation commits one to the view that no complete specification of standards of rationality can be given. This, however, does not frustrate the programme of giving a full-blooded theory of meaning for a language since the aim of such a theory is to give an account of what it is to speak a language and it is consistent with that aim to ground language in more general rational practices which are not described in the theory itself.

Consider this example of understanding the concept *square* which

McDowell takes from Dummett's (1978b). There Dummett notes that to have the concept square is to be able to discriminate between square things and things of other sorts. One of the ways of doing this is to apply the word "square" to recognizably square things. Two points should be noted. First, one could, prelinguistically, have the concept square but having the concept is not a *precondition* for learning the meaning of "square". Rather learning the correct use of the word "square" is one way of grasping the concept square. The point here is that we do not use the idea of a prelinguistic grasp of the concept to explain how meaning is imbued in language, i.e., how symbols come to have a representational rôle, rather, the possibility of a prelinguistic grasp of the concept entitles us to claim that the ability we are ascribing does not depend upon any specifically linguistic achievement. One could say that language is meaningful only because behaviour is. The theory of meaning gives an account only of linguistic meaning. Secondly, what one manifests in grasping the concept *square* is not a regularity in behaving differently relative to square things but a *discriminatory* ability, that is, a sensitivity to success or failure in this practice. This sensitivty to correctness (i.e., this element of rationality) is not essentially linguistic so, although fundamental to meaningful use of language need not be part of a complete description of what it is to use a language. Thus we can give an account of language which grounds possession of linguistically expressible concepts in non-linguistic but rational practices. Moreover, in doing so, we do not explain meaningful use in terms of antecedently grasped concepts.

This concern with the distinction between a regularity in behaviour and behaviour or use which is subject to standards of correctness readily calls to mind Wittgenstein's preoccupation with what it is to follow a rule in *Philosophical Investigations* and *Remarks on the Foundations of Mathematics*. McDowell's reservations about full-bloodedness seem to stem from his reading of Wittgenstein. In

Wittgenstein on Following a Rule (McDowell 1984) he argues that Wittgenstein's argument militates against two misconceptions of meaning; i) that grasp of meaning always involves an interpretation and, ii) that grasp of meaning establishes a sublime mechanism which determines use. The second alternative is simply a myth and the first delivers the sceptical paradox that meaning is impossible. The solution is to acknowledge that there is a way of grasping a rule which is not an interpretation and that this involves *acting within a practice*. One philosophical consequence of this is that our search for explanations eventually must come to an end. We hit "bedrock" and are reduced to saying simply, "This is what we do." The hard question is gauging just when we have hit "bedrock". Commenting on Wittgenstein's pronouncement that to use an expression without justification is not to use it without right McDowell writes,

> ...it seems clear that the point of this is precisely to prevent the leaching out of norms from our picture of "bedrock" -from our picture, that is, of how things are at the deepest level at which we may sensibly contemplate the place of language in the world. (McDowell 1984, p. 341)

I agree with McDowell's interpretation here and, if this is the import of his modest outlook, then must admit to sharing it. But I take Dummett's introduction of the idea that the manifestability constraint involves description of rational behaviour precisely to accommodate this picture. We ground language in practices which themselves involve normative constraints but that does not compromise Dummett's notion of full-bloodedness as I interpret it. That I am in danger of having reduced the point of contention between full-bloodedness and modesty to vacuity seems to be confirmed by the following quote from McDowell, which could be interpreted as favouring the notion of full-bloodedness that I ascribe to Dummett,

> Understanding linguistic behaviour, and hence understanding languages, involves no more than a special case of what understanding behaviour, in general involves. (McDowell 1976, p.45)

But my intention here has not been to side either with Dummett or with McDowell. I hope to have used the writings of both men to clarify what we can expect and demand of a theory of meaning whether one calls it full-blooded or modest. (I shall return to some of the issues involved in this question of normative control of a practice in connection with Crispin Wright's exploration of the consequences of Wittgenstein's Rule Following Considerations.)

§5 A Challenge for a Theory of Meaning:

For two classes of case giving a contentful account of that in which an understanding of sentences in either class consists proves relatively unproblematic. However these classes do not exhaust our linguistic repertoire. The first class consists of those sentences for which we have a suitably non-circular and explicit verbal explanation. Understanding of these sentences can thus be characterized as an explicit knowledge of this explanation. But, as I have noted above, this, on pain of circularity or regress, cannot provide the full story of our linguistic competence.

An important task of a theory of meaning will, thus, be to determine when it makes sense to ascribe implicit knowledge to a speaker. That is, where our theory is forced to characterize knowledge as implicit it must tell us when we are justified in making such attributions. We must be told what manifestable capacity constitutes knowledge of what our theoretical representation expresses. Our second unproblematic class of sentences requires us to make attributions of implicit knowledge. The sentences in this class are decidable so we can simply take the implicit knowledge to consist in the ability to exercise the appropriate recognitional capacity which, by hypothesis, we have or can acquire.

Non-effectively decidable statements - statements for which we have no method guaranteeing a decision on truth-value - pose a problem just because the ability to give an explicit explanation cannot always manifest understanding and it is unclear what ascriptions of implicit knowledge here amount to. We cannot always give a verbal explanation of these statements in the sense of being able to give a reductio*nist* account. The most we can expect is a reduct*ive* account, i.e., we can claim that the truth of statements in the problem class depends upon that of statements in the unproblematic class without presupposing a translation of sentences in one into sentences in the other, (See Dummett 1982 p.70 for this distinction.) This is because from a stock of decidable sentences we can generate only more complex but still decidable sentences unless we introduce an operator (such as quantification over infinite domains, tense operators and subjunctive conditionalisation) "responsible" for undecidability, i.e., an operator which, acting on decidable sentences, can produce undecidable sentences. The meaning of such an operator needs to be explained and cannot be given in terms of our basic stock of decidable sentences. (Just because in any putative explanation the property of decidability would be preserved.)

An attribution of implicit knowledge must, as noted, be justified by saying what it is to exercise that knowledge in demonstrable ways. if a speaker is said to possess implicit knowledge we must know what is being ascribed to him. We can only make contentful acriptions of knowledge where the would-be knower can distinctively manifest his possession of that knowledge. Dummett notes that,

> [i]mplicit knowledge cannot, however, meaningfully be ascribed to someone unless it is possible to say in what manifestation of that knowledge consists. (1975b, p.217)

So implicit knowledge must be specified as the ability to perform appropriately in circumstances recognizable by the speaker. If the circumstances adumbrated elude the speaker's ability to tell whether they obtain or not then it would never be possible to substantiate an accusation of ignorance since failure to respond appropriately will always be explicable as an inability to recognize that circumstances are, in fact, thus and so. It would not be possible to make a distinction between having and not having the implicit knowledge and so there would be no sense in making ascriptions of such knowledge. This point is not verificationist, it is a metaphysical or constitutive feature of our concept of knowledge. The point could be put as a challenge to a realist who claims that there is content to ascriptions of knowledge where no manifestation of possession of the knowledge is possible. It is then simply utterly obscure what explanatory value this notion of knowledge might have, given that it discerns no difference between those who possess and those who lack such knowledge. Or, alternatively, it might be agreed that realism was a stable position, so that being realist about knowledge ascriptions would support a realism about the truth of statements more generally, but the question would still remain as to whether there is a non-circular justification of realism.

Note that the position described is, in one sense, non-revisionary: we do not question whether or not we understand non-effectively decidable sentences. Our grasp of sentences of the first order practice is taken as a datum. What we then question is what this grasp consists in. A semantic theory which offers an account of the meaning of these sentences is then appraised according to its ability to issue in an account of our knowledge of that meaning which satisfies the manifestation and acquisition challenges. (This point will be important in the discussion in chapter 4 of the nature of a strict finitist position which is, supposedly, motivated by general anti-realist arguments. It is evident that such a position will not question whether or not we grasp certain sentences of the first order practice.)

In summary, we have the following broad perspective. The attempt to appraise a practice philosophically is taken to consist in an inquiry into the form which an appropriate theory of meaning for that practice would take. The theory of meaning must allow its possessor non-circular access to the meaning of every expression in the relevant fragment of language and, in doing so, must not assume grasp of a set of primitive expressions. The theory of meaning gives a theory of understanding, that is, it characterizes what is known by a speaker when he understands the language. Thus the theory is led to make certain ascriptions of knowledge to speakers. In making such ascriptions of knowledge the theory must give an account of what possession of that knowledge consists in; it must show how possession of that knowledge is both manifestable and acquirable.

§6 The Justification of Deduction:

Our attempt to justify a *practice* (as opposed to statements made within the practice; actual truth-values of and evidence for individual statements are not important here) consists in trying to give a non-circular account of the meaning of each expression in the practice. It is not obvious that this need always be possible nor that its impossibility should impugn the meaningfulness of the practice. Inferential practice, *prima facie*, offers a good example of this. Deduction is heavily enmeshed in our use of language as a whole. We use inferences to make appropriate passages between sentences, to work out a nexus of commitment associated with a given sentence. These commitments seem on the one hand to be constitutive of and, on the other, to be consequences of meaning. Clearly, inferential relations between sentences are intimately related to any plausible notion that we may have of the content of sentences. We often make deductions to glean information or to carry out a decision procedure. To deny the meaningfulness of inference (i.e., to accept inferential rules simply as operational procedures with no semantic contribution) is tantamount to denying the meaningfulness of language as a whole. So it seems we must include inferential practice in our semantic description of language.

Notwithstanding the force of that reflection there seems to be a powerful argument for denying the possibility of giving a justification of deduction. Any justification we are able to give will doubtless itself use inference. Those inferential practices themselves stand in need of justification so either the account is circular or it goes into a regress. Or, put another way, if we are able to deduce a rule of inference from certain others then we take this to be its justification but leave off regarding it as fundamental. There must be a non-empty set of fundamental rules which cannot be justified in this way. So either we justify them in terms of another system which launches a regress or else they are taken to be self-justifying which is circular.

Dummett tries to show (in *The Justification of Deduction* (1975a)) that this argument appears compelling only until we have distinguished two sorts of explanation and noticed the precise character of the accusation of circularity. Turning to the latter first, the putative justification is accused of being circular not because it assumes certain inferential rules as *premisses* but because the explanation *uses* those rules of inference. This can be viciously circular or not depending on the purpose of the explanation. Were it our intention to convince someone to use our inferential rules, i.e., to give a suasive justification, then the explanation would obviously be useless -receipt of the force of the explanation would presuppose acceptance of the rules of inference. This however is clearly not our intention. We want an account of the rôle of the presently accepted inferential practice in our language, i.e., we want an explanatory justification. So we can presuppose an acceptance of our inferential rules and offer a justification of them in terms, say, of a soundness proof.

Consider the following (attempted) justification of classical logic. Realism takes truth to be a property which a statement either has or does not have independently of our ability to determine the matter. Adherence to such a notion of truth is manifested by holding that bivalence is applicable to all sentences of the language. Thus, assuming the legitimacy of realist truth, we allow a justification of the classical logician's use of the Law of Excluded Middle (LEM) or, equivalently, of Double Negation Elimination (DNE) since affirming bivalence in this context is tantamount to affirming the Law of Excluded Middle (since asserting *P* is equivalent to asserting that "*P*" is true). So classical logic is given a semantic justification in terms of transferring truth (as realistically conceived) from premisses to conclusion. This only succeeds as a justification if an account can be given of what an understanding of (rather than a commitment to) realist truth consists in.

To tackle this question, consider a base class of sentences (which are both decidable and non-effectively decidable -since we may be interested in statements about, say, the past which may, presently be undecidable yet are apparently logically simple) which contains no logical machinery. Understanding, as characterized here, will be exhausted by an ability to use the sentences correctly, i.e., to know for each sentence what its conditions of correct assertiblity are. These conditions, since necessarily circumscribed by uses we are capable of making, must be conditions we can effectively recognize. (Note that there is no inference from this observation to the *denial* that we grasp non-effectively decidable sentences in the base class.)

The introduction of classical logic allows us to assert for any sentence, *P*, *P* or not-*P*, where, given the classical truth functional definition of the connectives, this assertion is the same as the assertion that *P* is either true or false (again, independently of whether we can determine the matter). Clearly, mastery of this notion of truth need not have been involved in learning the meaning of any sentence in the base class just because that mastery can only make use of our ability to recognize certain conditions as obtaining or not. So comprehending this notion of truth seems to be exhausted just in an acceptance of classically valid rules of inference. In other words we are forced to ascribe to ourselves knowledge of realist truth purely to validate classical rules of inference. There is a straightforward implausibility (if not an incoherence) here. It seems utterly mysterious to view this knowledge as constituted by acceptance of these rules of inference. Moreover it appears circular to seek to justify the logic by appeal to a concept which is manifested *only* in the acceptance of that logic.

In the *Justification of Deduction* (1975a) Dummett has two main concerns. First, he wants to show how a justification of deduction is possible. He finds that it is only from a certain conception of the nature of meaning, that is, from a molecular conception of meaning (in which grasp each expression can be given a non-circular explanation independently a grasp of the entire language) that a justification is possible. Holism, it emerges, rather than offering an alternative form of justification is the denial that a justification is possible at all. So a justification of deduction is appraised not by the logical theory but by the theory of meaning for the language, that is, we test the adequacy of a justification of deduction by asking whether the central concept used in giving a soundness proof for the logical system coheres with a molecular account of language. The second of Dummett's concerns now arises for it is a test of the adequacy of a theory of meaning that it explain how inference is possible. This, Dummett points out is no easy task because there seems to be a tension between how we account for the legitimacy of deduction and how we account for its usefulness. Classically, a deduction is legitimate whenever it is impossible for the premisses to be true and the conclusion false. It is tempting to think that this means that in acknowledging the truth of the premisses we have, in some sense, implicitly acknowledged the truth of the conclusion. This thought must be true to at least this extent, aside from acknowledging the soundness of the individual steps of the deduction no extra epistemic step is necessary so that, in the particular case of an immediate (one step) consequence the only epistemic advance we make is in acknowledging the truth of the consequence which is just to accept the soundness of the inference. The point now is that in order to account for the latter we want to say that the former is, in some sense, no significant achievement. But we also want to say that deduction is useful just because it represents some epistemic advance so that the achievement we have just tried to downgrade in accounting for the soundness of inference must be of some significance. The central problem with the first concern is in accounting for the use of the putative justification *as a justification* despite its circularity. To resolve the tension of the second concern we must make some concession to realism, admit some gap between truth and its recognition (at least, by direct or canonical means) but we continue to object to realist use of an epistemically unconstrained notion of truth for making impossible an account of what it is to understand and learn a language. An intuitionistic semantics thus receives tentative, programmatic support.

Susan Haack (Haack 1982) contests Dummett's claim that circularity is fatal only if our explanation is intended to be suasive. Haack's argument has two strands. She disagrees with Dummett that there is a disanalogy between an inductive justification of induction and a deductive justification of deduction. Also she argues that allowing circular justifications enables us to justify too much and so the justification becomes worthless. The first strand of the argument can be neglected here; the disanalogy between the two sorts of justification is not central to Dummett's claim that in the case of deduction such a justification serves our interests. The value of an inductive justification of induction can be left open to question. The second argument is obviously more powerful since it directly attacks the value of the circular, explanatory justification.

Haack points out that a circular justification is available even for such patently invalid rules as Modus Morons (MM), i.e., from $P \to Q$ and Q infer P. If we regard MM as valid then we can show that MM is valid. Suppose $P \to Q$ and Q are true. By the truth table definition of $P \to Q$, if P is true then Q is true. Also, Q is true. So, by MM, P is true. To object that this "justification" fails because MM is invalid is to beg the question. However it is too swift to conclude from this that the circular justification is valueless since if we have independent reason for believing that MM is invalid then the justification must surely fail since then we put ourselves in the position of requiring a suasive justification. We do have just such a reason because it is easy to show that using MM leads to a contradiction. Haack however objects that this demonstration assumes that the system in which we embed MM otherwise functions as "normal". The thought here is presumably not the simple one that changing the other rules of the system will return consistency. In that case we no longer have an obvious interpretation for MM and so our motive for regarding it as invalid is undercut. Indeed if we interpret "$P \to Q$" as P if Q then MM will be valid. Rather it seems Haack is suggesting that consistency (and conversely, contradiction) is a property of the system itself, of the entire set of rules not of any individual rule. Clearly contradiction arises as a result of what we feel jointly committed to. There is however, as yet, no motivation for taking it to be a property of the system as a whole. We could distinguish certain rules as giving the meaning of the logical constants and then demand that other rules be faithful to that meaning. In that way we could require consistency via this process of harmonising of rules. So in the present case we could demand that MM as an elimination rule be faithful to the meaning of "$\to$" as set up by CP (ie. conditional proof; from $\Delta, P \vdash Q$ infer $\Delta \vdash P \to Q$). We could then easily show it to result in a contradiction (take $P \vdash P$ then CP gives $\vdash P \to P$ and so also $P \& \sim P \vdash P \to P$, thus, by CP again, $\vdash (P \& \sim P) \to (P \to P)$ and, finally, MM gives us $\vdash P \& \sim P$). These questions however depend for their resolution on more general notions about meaning which we shall return to below. No direct attack is mounted on the specific circularity of the account or, rather, the attack on circularity only succeeds when it conspires with a repudiation of the idea that we can establish a meaning for a logical constant by stipulating an introduction or an elimination rule and insisting on harmony.

Before considering the validity of the suasive/explanatory distinction I want to remark on Haack's argument on the lack of discrimination of circular justifications. First, if the suasive/explanatory distinction is viable then we cannot, in one sense, land up justifying too much. We cannot justify MM in this way because it is never a candidate to be so justified. We can only justify deductive inferences that we, in fact, accept. There may be a legitimate question why we don't accept MM but that is a distinct question. So conceived then, the justificatory project can justify at most our present practice. It cannot force us to adopt new forms of inference but can lead us only to renounce certain forms of inference. Quite how the project gains this revisionary authority despite its acceptance of circularity is a matter we shall return to below. The fact, though, that Dummett takes the project to have this force should serve to warn us that Haack's objections to the lack of discrimination of this form of justification betoken a neglect of further constraints that Dummett places on the acceptability of a justification. Haack's argument thus only shows that we need to consider Dummett's position as a whole. It then becomes apparent that Haack has not thus far revealed an incoherency in the position, rather she objects to the position from an external perspective sympathetic to holism. I'm not suggesting here that the objection is not worth facing but simply trying to place the impact of the objection.

§7 The Distinction between Suasive and Explanatory Justifications:

Returning now to the suasive/explanatory distinction: a suasive argument seeks to convince so it must work from accepted premisses to the desired conclusion. The logical and epistemic directions must coincide. In an explanatory argument we seek to explain something we already accept so the premisses themselves need not enjoy the same degree of acceptability in advance. Indeed, Dummett holds that we may only accept the premisses because they provide an explanation of an already accepted conclusion. So here the epistemological direction runs counter to the logical direction.

Haack objects to the distinction in the acceptability of an explanatory rather than a suasive justification because it is drawn on the basis of what we believe. Thus the asymmetry in using, say, MPP to justify MPP rather than using MM to justify MM is not the result of the former *being* valid but only in our *belief* that it is valid. But the interest of the programme is in determining whether our beliefs are justified. That goal itself motivates the basis for the distinction.

Insofar as we regard an accepted statement as in need of and so capable of justification we will regard a set of premisses as supported by the fact that they provide an explanation of our acceptance of that statement. Where the premisses provide the only possible explanation then that support will be as strong as our commitment to the original statement. (As long as we hold that all aspects of our practice must yield to some form of explanation.) Where we are faced with a choice in explanations then the matter will be decided by what independent reasons we have for accepting each set of premisses. In the present case Dummett argues that independent considerations about the nature of meaning determine which explanatory justification we may legitimately endorse. My purpose in bringing this point up here is not to embark on an analysis of the nature of explanation but simply to motivate

Dummett's thought that in non-suasive justifications epistemic and logical directions need not coincide. Of course that observation is far from condoning blatant circularity in an argument. We can see, though, that even on this simple picture overtly circular arguments (arguments in which the conclusion is one of the premisses of the argument) will only be treated as degenerate arguments in the sense that they will have the form of acceptable arguments but will never be preferred over alternative explanations. The reason for this is just that when faced with a choice we seek independent reason to accept the premisses. Such independent reason in the case of a circular argument will constitute an alternative explanation of the original conclusion and, if persuasive as a reason for accepting the circular explanation, will simultaneously undercut that support by providing a stronger explanation.

What difference does the fact make that in a deductive justification of deduction the circularity is not blatant? What does someone learn from a justification of, say, MPP that uses rather than assumes MPP? Consider the following arguments:

A_1: Suppose MPP is necessarily truth preserving. So MPP is necessarily truth preserving.

A_2: Suppose that $P \rightarrow Q$ is true and that P is true. By the truth-table definition of "$\rightarrow$", if $P \rightarrow Q$ is true and P is true then Q is true. So (by MPP or its metalinguistic equivalent) Q is true.

A_1 is obviously a valid argument that could be accepted even by someone who either rejects MPP or who fails to understand it. A_2 on the other hand is only acceptable if MPP is itself acceptable. The difference is precisely analogous to the one between knowing a proposition to be true and knowing what the proposition expresses which Dummett (Dummett 1975c, p.107) draws attention to. A_1 is entirely uninformative because to progress from knowing it to be sound to knowing what it expresses one needs to understand the very rule being justified. One cannot, though, even get as far as accepting A_2 to be sound unless one already accepts MPP. Recall that Dummett uses the distinction to show why the M-sentences of a Davidsonian truth theory are uninformative. The Davidsonian programme is thus faced with the difficulty of either having to accept that its axioms make the resulting theory modest (i.e., no account of what it is to know what they express is essayed) or holistic. Is A_2 compatible with the aims of a full-blooded theory of meaning or does full-bloodedness require the use only of suasive arguments?

There is a strong reason why we should regard A_2 as compatible with the demands of full-bloodedness and that is that a theory using A_2 will only be accepted if MPP itself is accepted so that, although A_2 uses MPP, we can accept it as giving an explanation of MPP (on the proviso that the logic of the metalanguage includes MPP). Analogously, provided the meta-language includes the word "square", we can give an account of the concept *square* by using "square". But could the theory be used to justify and, in this sense explain, MPP? Dummett (1987), in his reply to John McDowell's (1987), raises a similar problem for the classical logician in his communication of the law of excluded middle (LEM) to a doubting intuitionist. Despite the intuitionist's acceptance that the logical constants are subject to their usual truth-table definitions he contests the classical logician's use of LEM by asking with what right the classicist assumes that the two lines of the truth-table exhaust all the possibilities or, better, that one or other line must relate determinately to the actual state of affairs. That assumption, the assumption that one or other combination of truth values determinately obtains, is equivalent to assuming LEM so can't be used in justifying/communicating the latter to the intuitionist. Of course there is no contradiction here with Dummett's defence of a deductive justification of deduction since the position just outlined was one that called for a suasive argument. The question is whether Dummett's interests in the theory of meaning are only to be served by suasive arguments. I want now to consider in a little more detail the precise demands of full-bloodedness.

Later in the same paper Dummett says,

> ...a semantic theory that is maximally stable under changes of the underlying logic of the metalanguage imparts an understanding of the logical constants, as used in the given logic, to whoever accepts those laws under whose replacement it is not stable. (1987, p.267)

The virtue of this position Dummett claims is that we get as full an explanation of the logical constants as is possible. This statement, however, wants a gloss. The sort of justification just offered of LEM by the classicist is maximally unstable since it succeeds in justifying just those laws which are accepted in the metalanguage. Hence the intuitionist's objection: if he insists that the semantic theory be given only in a metalanguage using intuitionistic logic then he undercuts the classicist's justification of LEM. In contrast an intuitionistic semantics used in a metalanguage obeying classical logic will succeed in explaining to the classicist the intuitionistic demonstration of the invalidity of LEM in the object language. A prerequisite to resolving a dispute between parties disagreeing about the validity of certain logical laws is that mutual understanding is reached. But, in virtue of the fact that the dispute is about the validity of certain rules of inference, mutual comprehension is difficult to achieve because in explaining the meaning he attaches to the logical constants each logician will be using rules the other fails to accept. We can only achieve mutual understanding if a semantics is given which is appropriately insensitive to changes in the underlying logic of the metalanguage so that both logicians can take their logic to be operative in the metalanguage. This is only a prerequisite to resolving the dispute because the matter is not purely technical but also relies on philosophical considerations, general ideas about meaning, by which we test the notions used in the semantics.

This description suggests the following picture. Starting from a maximally unstable semantics we can progressively show the invalidity of certain laws. Thus we embark on a revisionary programme of reducing the power of the logic. At each stage the semantics given is more stable just because it is stable relative to substitution of the more powerful logics in the metalanguage. In this way we finally achieve a logic justified by a maximally stable semantic theory. But if this is the only constraint operating on the project why don't we succeed in abolishing inference by adopting an extreme constructivist position (one which only admits a notion of truth in which truth is identified with its actual recognition by direct means so that we only accept assertion of logically complex sentences when canonical conditions for use of the introduction rules are satisfied)? The thought here is that statements asserted as a result of using an elimination rule (e.g. assertion of the logically simple P via MPP on the already assertible Q and $Q \rightarrow P$) are not always known by direct or canonical means to be true. Thus either any remotely useful inference becomes invalid or we have to show that despite lack of recognition a canonical warrant always exists whenever we have an inferential warrant. The problem however is to demonstrate that the latter is so, given only the inferential techniques provided by the extreme constructivist. Since the inferential techniques available on this radical constructivist position are so meagre this amounts to having to give a suasive justification of deduction, i.e., justifying the inferential practice by showing, without use of inferential techniques, that the use of inference preserves assertibility with respect to canonical warrants.

Dummett uses the contrast between an idealist or constructivist and a realist theory of meaning in focusing on the tension in trying to account for both the legitimacy and the usefulness of deduction. The tension arises because we want to say that an inference is valid just because in recognizing the truth of the premisses we have, in a sense, accomplished all we need to do to recognize the truth of the conclusion, that that step of assent involves no further cognitive achievement. Alternatively, the state of information required to recognize the premisses as true suffices for recognition that the conclusion is true. On the other hand inference is useful just because it is informative, the practice of inference reveals that certain propositions hitherto unrecognized as true are, in fact, so. Haack complains that this is only a pseudo-problem arising from an equivocation in Dummett's use of deductive inference to mean both deductive implication (the logical relation that obtains between propositions) and deductive inference (the actual inferences drawn by a subject). The connection between these two notions is that x correctly infers Q from P iff P deductively implies Q. (The "only if" part of this biconditional should be clear. The "if" part is doubtful since we do not want the question of whether P deductively implies Q to depend on whether anyone actually draws the inference correctly. We can, however, insist that the deductive implication holds just when there is a possibility, however characterized, of drawing the inference. That possibility can then be seen as relating to a notion of a (possible) subject in appropriately ideal epistemic conditions, with ideal cognitive capacities and the desire to draw the inference correctly. Insisting on the "if" part of the biconditional is then a way of questioning how this notion of possibility should be characterized in terms of an ideal subject.) The resolution of the tension exploits the distinction as follows. P deductively implies Q just when if P is recognizable as true so is Q. Note that this only says that being in a position to recognize P as true is to be in a position in which Q is recognizable (not recognized) as true. So the practice of inference is useful because it shows the cognitively limited logician that he is in fact in a position to recognize a given proposition as true. Haack insists that in order to recreate the tension we would have to incorporate either a logical element in the psychological account (a weak strategy) or a psychological element in the logical account.

For a proposition to be recognizable (as opposed to recognized) as true is for us to have an effective means which allows us, at least in principle, to recognize that the proposition is true. This gives rise to the distinction between coming to recognize a proposition as true by direct and by indirect means since if we are capable of recognizing that we have such an effective procedure then we are obliged to assent to the proposition. It is just this distinction which, Dummett argues, is necessary to reconcile the tension and which betokens a concession to realism since it allows that the truth of a proposition is not to be identified with recognition of its truth, at least, by direct or canonical means. So it is arguable that Haack only succeeds in dismissing Dummett's problem by assuming his solution. This argument, however, relied on identifying indirect verification with recognizability. But what if notions of recognizability can be explained somehow else or accepted anti-realistically as primitives?

Haack hopes to show that no concession to realism is required in order to account for the informativeness of deductive inference. Let us reconsider the argument. We have the equivalent statements:

"P deductively implies Q" (1)

and "x correctly infers Q from P". (2)

Now (1) is unpacked as:

"when P is recognizable as true so is Q" (3).

So, by (3), no further epistemic progress is involved in recognizing Q as true once P has been so recognized since otherwise having recognized the truth of P we would need some further acquaintance with the world in order to recognize the truth of Q. If this further matter of acquaintance were a necessary fact then we are in danger of developing a regress (in order to know that P implies Q we need first to know whether P implies R, say) whilst if it were a contingent fact then we have the possibility of (3) being counterexemplified. Haack introduces the possibility of informativeness of inference by noticing that the use of recognizability in (3) allows fallibility on the part of the subject i.e. (3) is consistent with:

"x recognizes P as true but does not recognize Q as true" (4)

This makes sense of the "correctly" in (2), showing the normative role of logical propositions. But now the logical form of (4) is slightly ambiguous since what it does is point to a possibility which is compatible with (3) and it is not clear how we should characterize this possibility. I am not certain what interpretation Haack favours. She notes that "[i]f the premisses of a logically valid argument are assertible, so too is the conclusion; but human cognitive limitations are such that we may assert the premisses yet fail to assert the conclusion" (1982, p.233). The sentence is obscure. If, in the circumstances described, failure to assert the conclusion is a cognitive failing then the informativeness of deduction is already presupposed, for that position was just the one of holding that a genuine cognitive achievement was involved in acknowledging the assertibility of the conclusion. Ignoring, therefore, her use of the adjective, does Haack intend us to read: "Human limitations are such that we make mistakes so we may assert the premisses ..." (call this interpretation (4')) or "Human limitations are such that the facts of the matter are not apt for human recognition so we may assert the premisses ..." (call this interpretation (4")))? Haack gives no motivation for the latter, stronger reading.

(4') is entirely uncontroversial. It is obscure how such a banal observation is to enable us to account for the informativeness of deduction. It is banal because all it allows is that there may be instances of fallibility in inference so that it is, in a sense, informative to each of us to get his inferences right, to reflect correctly on his state of information. But that is not the question at issue, we want to know how inference can be useful to us (communally). The contrast is that between someone who knows that P and knows that Q and so recognizes from within the practice of inference that he also knows the proposition P-and-Q (so, in a sense, learns something), and how the practice itself can be informative or useful, that is, what purpose conjunction introduction has in the language.

The acceptability of (4") to a constructivist is, at best, contentious so Haack's claim that the apparent need for a concession to realism arises out of a mere equivocation grossly simplifies the situation and, in doing so, loses the point at issue. The constructivist introduces his semantics as a response to a realist semantics where that semantics is objectionable because of its use of a notion of truth which determinately applies or fails to apply although we can have no guarantee of being able to recognize it as obtaining whenever it does so. (4") combined with (3) ascribes the self-same property to recognizability since it says that a proposition may be determinately (this from (3)) recognizable as true although it is never so recognized and where the feat of recognition may be one that is guaranteed to evade us. So if we accept this as Haack's intended interpretation of (4) then it is entirely unsurprising that in this way she shows that no concession to realism is called for since the concession to realism has already been made in the setting up of the semantics. The shibboleth of a realist position can be stated in terms of what is taken to constitute a possible world in which a given proposition is recognized as true, the realist maintaining that a world is still possible even if it includes beings with infinitary idealisations of our recognitional capacities. Here, there is no need to excavate the intricacies of that debate (that I shall attempt later in the chapter on strict finitism); it is enough to observe that Haack in failing to distinguish (4') and (4") and so noticing no obligation to argue for (4") has covered over an important issue which Dummett's discussion nicely highlights.

We can return now to Haack's suggestion that to see a tension in the possibility of deduction one needs to incorporate a psychological component in the logical account. Her rendition of how such an argument might go is as follows. Take logical necessity to be truth in virtue of meaning and meaning to be given in terms of warrants for assertion or conditions for recognition of truth. Then: "For a proposition to be logically necessary just is for it to be assertible, or recognizable as true, come what may. So if anyone recognizes the truth of the premisses of an argument but does not recognize the truth of the conclusion, then the connection between the truth of the premisses and the conclusion cannot be necessary" (1982, p.228). The argument clearly takes the constructivist to be denying fallibility. But it is a mistake to take the constructivist to be abolishing the distinction between inferring and inferring correctly. The constructivist simply insists that we give content to the latter only by appeal to actual capacities to use the language. That position enables the constructivist to make some link between deductive implication and our practice of inferring; enough of a link to outlaw a position such as that described in (4"), at least under certain interpretations of it. It is tendentious to describe this position as one of building a psychological component into the logical account for that is to ascribe to the project a spurious psychologism. The project is to avoid psychologism whilst insisting that in the account of deductive implication we pay due heed to general considerations about meaning, and that means the account is constrained by reference to human recognitional capacities.

Haack consistently fragments Dummett's project, attacking each element in isolation without considering the way other elements set up constraints on solutions of a particular aspect. Thus for Dummett circular justifications are constrained by the form of the theory of meaning and deductive implication is connected to inference via the theory of meaning. In view of the central role that the theory of meaning has in Dummett's philosophy it seems unfair to complain that Haack's disagreement with Dummett is just a difference about the nature of meaning and the rôle of the theory of meaning in philosophy. But Haack's arguments were supposed to counter Dummett's views on the matter and to support her own. Far from doing that, her arguments are premissed on the acceptability of her own position, they receive no independent motivation and, symmetrically, she motivates no departure from Dummett's position.

The validity of Haack's criticisms was only an ulterior motive in this discussion. The main interest was in reflecting on why we are not driven into the extreme constructivst position in our search for a maximally stable logic. In support of Dummett we found that there was a counter-weight to this programme in the need to give an account of the usefulness of deduction. (This consideration will play an important role in my discussion of strict finitism.) This does indeed rely on a liberalization of the notion of truth beyond that which ties truth to recognized truth. But we shall see that truth as so conceived is not anti-realistically objectionable.

§8 What does Dummett take to be the revisionary implications of these reflections?

Classical logic can only be justified by appeal to a notion of truth which determinately applies or fails to apply independently of our abilities to determine the matter. Since such a concept is repugnant to our view of the nature of meaning we cannot justify classically accepted inferential practice. That practice, therefore, stands in need of revision.

We have had a convincing argument for the second premise. What of the first? The question is large. Dummett's view on the matter is that local justification of classical inferential practice is possible since many local anti-realisms have just this form. An example of this (see Dummett 1981, p.437) is neutralism about the future, that is, the belief that the future consists of a set of definite future courses of events. A statement is true if true in every such course and false if it is not.

Thus, on one interpretation, every statement is either true or false and classical logic applies. Yet the position is not realist since the semantics is not a classical two-valued semantics, the notion of truth is given a reductive account in terms of true-in-a-future-course-of-events, we do not suppose that there is a future which determinately makes our statements true or false.

The most obvious tactic for justifying classical logic is via the notion of assertion. However Dummett shows that using the principle that assertions cannot be neither correct nor incorrect only reinstates the principle that assertions cannot be neither true nor false (i.e., tertium non datur) and it takes classical logic to convert this to the principle of bivalence. (Briefly the argument takes something like this form. Consider the following argument for Bivalence. Intuitionism claims that we do have, and so tries to give an account of, an understanding of undecidable predicates. Now truth that transcends verification is just such a predicate and presumably the intuitionist will give an account of this understanding as he will for other undecidable predicates. If he does this then we must know of any sentence what it is for it to be true. So we can identify the meaning of each sentence with its truth conditions. If we know what a sentence's truth conditions are then we may not be able to tell whether or not in any particular case they are fulfilled but we can acknowledge only two possibilities, since otherwise we threaten the idea that we grasp the truth conditions. This last step follows by the intuitionist's own requirements for the intuitionist holds that if we have an understanding of a concept then if we are not able, in principle, to determine that it is applicable it is not applicable. To be more precise, the intuitionistic account of negation makes the negation of a sentence assertible just when we could draw a contradiction from the assertibility of the sentence itself. Thus if we can show that, in a suitable sense, a sentence is, in principle, not assertible we have a contradiction with the supposed assertibility of the sentence. So the negation of the sentence becomes assertible. Note, that the "in principle" qualification is important since without it we cannot suppose that the supposed assertibility of the sentence develops a *contradiction* as opposed to an implausibility or physical impossibility. So if it happens that a sentence is undecidable we cannot say both that we cannot determine whether its truth conditions hold and that they may hold. So either we have a means of determining the truth value or the sentence's truth conditions fail. Both positions allow us to assert bivalence. The last step of this however relies on the inference: $\sim(\sim P \& Q) \rightarrow (P \vee \sim Q)$. This is intuitionistically invalid so we need a realistic assumption to complete the argument.) For the present I want to note that Dummett's views do have revisionary implications in at least this sense. Accepting Dummett's methodological stance leads inexorably to repudiating an orthodox (realist) semantics given by truth-conditions. So it demands the construction of an alternative semantics. Whether an anti-realistically acceptable semantics is possible and what revision it would entail are questions I shall examine in relation to mathematics.

What form might such a semantics take? The central problem for realist truth is that it is unconstrained by our recognitional capacities. The gap between truth and our recognition of truth is filled only, if at all, by the notion that an assertion is true if it could be observed to be so by some ideal being with (perhaps) infinitary recognitional capacities. The conclusion of the anti-realist arguments is that thinking that we have this notion is a radical misconception of how we imbue language with meaning. Thus truth must be linked to our capacities to recognize, in principle, whether it obtains or not.

The picture encouraged by this description is that we identify truth with the notion taken as central in the semantic theory which gives an account of the content of assertions. (And so, in some attenuated sense, give a truth conditional account of meaning.) I take it that this is not the programme. We give the meaning of a sentence by specifying conditions which we can recognize either as obtaining or not obtaining and which warrant the assertion of the sentence. The conditions, since they warrant the assertion of the sentence, warrant the assertion that it is true. We cannot, however, take truth to be equivalent to these conditions themselves. The reason for this is that truth is an atemporal concept in the sense that revising a truth-value necessarily impugns a previous judgment. Another reason is that truth and assertibility conditions of a sentence may need to be distinguished when considering the sentence as a constituent of more complex sentences. (See chapter 4 for my argument that an account of disjunctive statements shows truth need not coincide with warranted assertibility.) However since identifying assertion conditions with truth conditions would entail that truth was always be a decidable property we must accept a distinction between our concept of warranted assertion and truth. We acknowledge this by recognizing that an assertion may be warranted yet false and not warranted yet true. So we may revise our opinion about the truth value of a statement without revising the question of whether or not we possessed a warrant. Warranted assertibility is a notion which recognizes the particular epistemic postion of a speaker. Truth, however, even if epistemically constrained relates to suitably ideal epistemic conditions. Our theory of meaning always gives a reductive account of truth either by characterizing meaning in terms of verification and/or falsification conditions, or conditions of justification or criterial warrants and then allowing an account of truth by utilising the equivalence principle. In other words, one characterizes the notion of truth through our practice of assertion. Or one may give directly a reductive account of truth and then characterize meaning in the style of Davidson by assuming a notion of truth in a Tarskian truth definition to deliver an account of meaning. But this second strategy, it should be noted, will not give a global account of meaning since we always require a base class of sentences in terms of which the reductive account of truth is to be given. This reflection might seem to have little bearing on the mathematical case where conditions of assertion are anyway indefeaible in that revision of a truth value assignment is impossible without impugning the previous claim that conditions warranting an assertion obtained. The dialectical position should however be made clear and may be important in the way we account for the introduction of informal proofs and new proof procedures. An intuitionistic semantics usually characterizes meaning in terms of some (recursively characterized) notion of canonical proof. We will (almost certainly) not want to identify truth in mathematics with possession of a canonical proof but will give a reductive account of truth in terms of canonical proofs. This would allow the possibility of us broadening our concept of proof to include informal proofs and to account for informal proofs as being necessarily truth preserving. Also if the notion of canonical proof is firmly circumscribed (by, say, being tied to a given formalisation) it would become difficult to explain the development of new proof procedures which transcend this formalisation unless the notion of truth transcended (even if explained in terms of) the existence of canonical proofs. (But see below -chapter 6- for my discussion of this issue.)

Dummett takes the notion of conditions (or a subset of conditions) warranting assertion as central in the construction of a theory of meaning for the language. There are good reasons for doing so. One reason is that this seems to preserve important insights of the truth-conditional account. That sort of semantics was appealing because the notion of truth relates precisely to the correctness and incorrectness of an assertion. An assertion is correct just when it is true. So, taking truth as a semantic primitive, we can spell out both the content and the aim of an assertion by reference to truth-conditions. We have forsworn use of realist truth as a semantic primitive but still take the practice of making assertions as fundamental, giving the content of assertions in terms of (defeasible) warrants for assertion. These in turn give an account of our use of the truth predicate in the language. The point of this is that, as well as utilising assertion for its rôle as a fundamental linguistic act, we develop an account of our concept of truth for the discourse and the metaphysical issue of realism relates precisely to this notion of truth.

A mathematical assertion is made in either of two cases, the possession of a proof for it or of an effective means guaranteeing construction of a proof. Thus in this case assertibility conditions relate directly to our ability to prove the sentence. We understand the sentence when we can recognize a proof of it. So we take a notion of provability as central in giving the meaning of mathematical sentences.

How do we characterize the meanings of the logical constants for mathematics? A natural first thought is that the meanings of the logical constants are given in the rules of inference. Given a set of (well-understood) logically simple sentences an understanding of the connectives seems to be exhausted by an ability to recognize and perform patterns of inference sanctioned by the canon of rules and to know that these patterns warrant the assertion of the concluding sentence (whether this is simple or complex). The thought here is that the rules of classical logic are constitutive of the meanings of the logical connectives themselves so that provided they do not lead us into contradiction, i.e., provided they give rise to a consistent practice, we should accept that the meanings of the constants are implicitly given in the practice itself. We simply need to be able to determine what counts as adherence to and what counts as contravention of the rules of inference and what rôle the practice of inference has in providing warrants for assertion. Described in this way the canon of rules must be adequate solely on the proviso that it can be mastered and that it does not lead to absurdity. Classical logic would then receive a rapid reinstatement.

To accept the existence (or possibility) of a practice is however not to condone it. The practice must be seen to have some point. Our justification should explain the purpose of the practice. Any inferential practice allows us to pass from one set of statements to others in accord with certain rules. That purely syntactic description leaves out an important element in the practice since it gives no clue as to what pedigree we take to be transferred in inference. As soon as we take on that obligation we incur the added responsibility of saying just why those inferential rules transfer that pedigree. That is, we must give an explanatory justification of deduction.

A sound inference is such that being in a position to assert the premises of the inference is just to be in a position to assert its conclusion. The inference transfers the property of being correctly assertible. So far that says very little. It certainly imposes no adequacy conditions on the inferential practice because it just is in the nature of our use of inference that it supplies a warrant for assertion: if the inference is correct then it is correct to assert its conclusion (if we are warranted in asserting its premisses) and it is correct just when it complies with the rules of inference. No adequacy condition is imposed just because of this circularity. The idea that inference is constrained by having to comply merely with the existence of a warrant for assertion is too crude to result in contentful constraints. Since inference itself gives us warrants for assertion no restrictions are encountered.

The only way this requirement can lead to adequacy constraints on inference is in the preservation of meanings of logically simple sentences given independently of inferential warrants. Alternatively, we insist that the possession of an inferential warrant must be constrained by the existence of (or our ability, in principle, to furnish) non-inferential warrants. Logically simple sentences which become assertible via inference should, independently of inference, be correctly assertible, i.e., the insertion of inferential rules into the base class should be a conservative extension of that class relative to correct assertibilty. These thoughts simply unpack the naïve view that inferential rules should be faithful to the meanings of sentences. Their importance is that they enable us to think of language as being partially ordered. The competence gained at one level is complete and survives intact the introduction of a new level of greater (logical) complexity. Thus we can account for our progressive acquisition of language. For any sentence only a proper fragment of language need be known for an understanding of that sentence.

The implication for the form of justification of logic is that we shall have to give a uniform account of meaning for both logical and non-logical vocabulary. Above I mentioned that we need to justify the inferential rules in terms of their ability to ensure that a certain pedigree is transferred in a sound inference from premisses to conclusion. The pedigree we are interested in, obviously, is truth, but the important point is to uncover what notion of truth we are entitled to credit ourselves with, given the nature of our understanding of logically simple sentences. We insist on uniformity in our account of logical and non-logical vocabulary by refusing to grant that we have grasp of a notion of truth which transcends that which we are forced to admit in our account of logically simple sentences. So if a logic can only be given a putative justification by appeal to such a notion of truth we repudiate that logic as unjustifiable. Technically, we enforce this constraint by insisting that our logic establish a conservative extension of the base class of sentences, that is, we insist that no sentence in the base class becomes assertible on the basis of an inferentially provided warrant if it was not, in some sense, assertible on the basis of a non-inferential warrant. Thus it will be insufficient to advert, as above, to the meaningfulness of the base class and non-collusive agreement in inferential practice. Those observations are preconditions for there being something to justify but don't amount in themselves to a justification. They give us no insight into the nature of the practice. Dummett claims that to accept this position is to accept holism.

§9 Summary:

In this chapter I have attempted to expose Dummett's arguments for taking the theory of meaning as a vital tool in our attempts to reach a philosophical understanding of a practice. I elucidated certain constraints which an acceptable account of meaning must satisfy if it is not to contravene basic requirements of a notion of meaning in showing how we learn a language and how we communicate through language. The result was that a theory of meaning had to meet the challenge of giving an account of just what our understanding of any expression consists in, in terms which show how, when a speaker is said to possess knowledge, that state of cognition is manifestable. This cannot be done if the theory makes use of notions, such as the realist or bivalent conception of truth, which are not linked to recognitional capacities. The theory of meaning can be seen to be full-blooded provided that requirement is suitably interpreted as not requiring a reduction of meaning to norm-free behaviour and/or dispositions. The possibility of giving a justification of deduction was argued for and objections raised by Haack to the circularity of that project were dismissed. A tension between the search for a "maximally stable" logic and an account of the usefulness of deduction was noted. (That issue is pursued in connection with the business of chapter 4 below.) Lastly, the need to give an account of the validity of deduction by showing how inference transfers some semantic property which is appropriately constrained epistemically shows that a classical two-valued semantic theory cannot be adhered to universally. So revision of logic is possible.

CHAPTER THREE: THE NEED FOR A JUSTIFICATION OF DEDUCTION

The Need for a Justification of Deduction

The scope of this chapter is limited. I attempt to rebut certain arguments for the conclusion that the anti-realist can and, perhaps, should eschew a justification of deduction. These arguments (gleaned mainly from the writings of Crispin Wright) have two main focusses. First, it is claimed that a semantic justification is not mandatory given only the anti-realist's constraint of molecularity. Secondly (and more powerfully) it is argued that such justifications presuppose a conception of meaning which is unavailable to the anti-realist since in the process of justification one assumes that there are determinate and transcendent conditions of correct use. My contrary claim is that a refusal to concern oneself with a semantic justification of deduction involves a departure from a molecular conception of meaning and that we have been given no reason for supposing that the justificatory project necessitates an anti-realistically repugnant notion of meaning. Thus the (molecular) anti-realist position is, at least, potentially revisionary since it requires a significant justification of the practice. The validity of classical logic will thus depend on an appropriate semantics being forthcoming. I do not claim that no anti-realistically acceptable validation of classical logic is possible. I do, however, accept that the classical two valued truth-functional account of classical logic involves an inadmissable concession to realism (or, perhaps, should rather be identified with what it is to assume a realist position with respect to a given discourse). So even if we can be brought to accept classical logic as valid, we are, as anti-realists, necessarily involved in revision of, at least, the semantic theory underpinning classical logic.

I do not argue (in this chapter or anywhere else) that an anti-realist position *must* coincide with a version of intuitionism. I leave open the question of whether an anti-realistically acceptable semantics which succeeds in justifying classical logic is possible. That is a large question. (It would, for a start, involve a survey of various super-valuational semantics and other reductive semantic accounts.) The support I offer for intuitionism consists, first, in explaining the intuitionistic semantics as a plausible (but by no means an inevitable) outcome of anti-realist semantic considerations and, secondly, in defending that position from certain accusations of incoherence. (This, I undertake mainly in chapters 4 and 6.) So the strength of my argument for intuitionism as a philosophy of mathematics is limited to showing that it is anti-realistically acceptable, not that it is mandatory to an anti-realist.

The course of my argument has been dictated by a set of considerations rallied by Wright. The strucure of the chapter is thus somewhat convoluted. I hope that by anticipating my argument here I can make that structure clearer. I begin (§1) by outlining the relevant notions of holism and molecularism. In §2 I offer a direct rebuttal of the consistency of holding both Wright's implicit definitional view of logic, and a molecular view of language.

§3 investigates a somewhat different and, perhaps, deeper challenge to revisionary anti-realism. The broad challenge (which depends on a complex set of considerations to do with the process of validating a logic) is for the anti-realist to give content to the specifically *semantic* notions he regards himself as entitled to. The point of the challenge is that unless the anti-realist can make good an accusation against a semantic theory of using anti-realistically repugnant notions the characterization of the metaphysical issue of realism in semantic terms threatens to collapse. My defence of anti-realism hinges on, i) reservations about the detail of the argument and, ii) a recognition that although the possibility of raising the metaphysical issue of realism requires that there be some measure of agreement between parties to the dispute, this need not be a source of concern provided no metaphysical question is begged.

§4 considers an argument of Wright's which purports to show that the anti-realist is unable to concern himself with questions of soundness because an instance of error can only be recognized from *within* an inferential system. There is thus no recognizable objective fact which determines whether or not a given system (short of inconsistency) is sound. I attempt to increase the tension by arguing that despite the relativity Wright claims to have revealed in questions of soundness *any* practitioner of inference cannot accept that more than one system is sound.

§5 attempts to respond to this tension. First, I claim that Wright has only succeeded in demonstrating the undecidability of questions of unsoundness. So from the intuitionistic point of view it has been shown that any system is not unsound (which is not, by intuitionistic lights, to say that it is sound). Secondly, I notice that the accusation made by the intuitionistic anti-realist against classical logic is not that it is unsound, it is that it is not demonstrably sound. It is quite possible for an anti-realist accepting the first position to raise questions about soundness (rather than unsoundness) and thus to make the accusation levelled in the second position.

In this last section (§6) I argue that Wright's reservations about revisionary anti-realism stem wholly from his interpretation of Wittgenstein's Rule Following Considerations. I offer an alternative interpretation of those considerations (drawn from the writings of John McDowell) which I claim is consistent with anti-realism. So no internal incoherency in that position has been revealed.

•

§1 Holism and Molecularism:

We are concerned with the effect of holism on the form of the theory of meaning. Thus we can think of holism as the denial of the possibility of a molecular theory of meaning. A molecular theory holds that understanding any expression requires competence in at most a fragment of language and that we can give an account for every expression of what that competence consists in. There are three ways of denying this, so three forms of holism.

Holism$_1$:- some sentences do not have a determinate meaning.

Holism$_2$:- all sentences have determinate meanings but understanding some sentences requires competence in the entire language.

Holism$_3$:- all sentences have determinate meanings not dependent on an understanding of the entire language but some expressions are such that we cannot give an explanation (in the sense of a semantic justification as glossed in the last chapter) of what an understanding of them consists in.

I should here add a word to my use of the notion of determinate meaning. The manner in which I would choose to elucidate this notion would be that an expression has determinate meaning if its meaning is capable of being explained within an appropriate theory of meaning for the language.

Holism$_3$ tries to meet many of the molecular constraints that emerge from considering the nature of understanding but is shy of the strength of the molecular methodological stance (which insists that it be possible to give a non-circular account of the meaning of each sentence in terms which presuppose competence in at most a fragment of the language) because of its acceptance of the possibility of imposing a revision in a first order practice. It allows the possibility of a practice evading explanation within a theory of meaning without seeing this, in molecular fashion, as a condemnation of the meaningfulness of the practice. We could thus take it that it differs only terminologically from the thorough molecular position since both might agree on the sort of description we can give of the practice, i.e., they need not differ on the question of whether and what sort of justification we can give but only over the need or rôle of the justification. However that would be to loose a, possibly important, distinction. The adoption of a methodology reveals what one takes to be the nature of meaning. $Holism_3$ encapsulates a certain concept of meaning. It is important because it is of crucial interest to our project to discover whether or not a departure from molecularism need collapse into a holism of the first two sorts since, as I have just noted, $Holism_3$ allows us to accept many of the semantic constraints governing our account of the practice yet provides a possible means of defusing any revisionary implications that might follow therefrom.

In this chapter I shall be interested in assessing the feasibility of a position which is a version of $Holism_3$. I want to argue that this version of $Holism_3$ collapses into a holism of the other two sorts. ($Holism_3$ obviously collapses into $Holism_1$ if the expressions it excepts from having a meaning theoretic explanation include *sentences.* So, in part what I shall be arguing is that having this attitude to subsentential expressions has discernable effects at the sentential level.)

§2 The Relation Between Holism and the Justification of Deduction:

It would be difficult to undertake this task without considering precisely which expressions this version of $Holism_3$ excepts from having a meaning theoretic explanation. Wright (1981) moots a view of *logic* which is of a form of $Holism_3$ since, according to this view, we do not give a complete account of all logical vocabulary in meaning theoretic terms. Dummett argues that the only justification we have of the global use of classical logic depends on an objectionable notion of truth. So we are led inexorably to revision of logic with holism (1 and 2) the only saving position. Wright questions whether the anti-realist might not disavow the need for a semantical justification of logic (so avoiding the trap of realism) yet maintain a molecular view of language. The picture envisaged is just that understanding the logical constants accrues through immersion in the practice of inference. The logical constants are thus implicitly defined. The picture does not, at least at first sight, describe a holism (of the first two sorts) since the essence of the molecular position is that complete understanding can be gained in just a fragment of language (that competence being undisturbed in the passage to increased linguistic skill). This segmentation, Wright claims, might still be possible although each stage requires complete logical competence.

The sort of justification we required was one that showed the role of inferential practice in language as a whole. This, if we follow Wright, we are now reneging on. The requirement was a methodological one so there need be, as noted before, no substantive disagreement, both parties agreeing that the original methodology demands some kind of revision. Wright however urges that nothing substantial is forfeited in shifting the methodological stance drawn from the demands of molecularism (i.e., that that position was not well motivated) and further (for reasons we shall encounter below) that some shift is forced on us since the original stance is internally incoherent.

The molecular view holds, first, that we can give a representation of the contents of individual sentences and, secondly, that we can do this independently of a description of the entire language. Assuming that we have isolated a logically complete stage then we have certainly satisfied the latter constraint, i.e., *if* we can give a representation of the content of individual sentences then that representation will be independent of a description of the entire language. Can we satisfy the antecedent of the conditional? What representation of the contents of individual sentences might we try to give? If we cannot give such representations then our position will collapse into Holism_1.

We might try to give the meaning of a sentence in terms of its position within the inferentially connected web of sentences or else we might try to isolate some feature of use relative to a specific subset of other sentences. The first position is as bad as requiring competence in the entire language because it implausibly makes understanding a sentence reliant on grasp of all the inferential connections in the language. This is a bad position partly because competence in a language never requires this sort of logical omniscience (in fact this omniscience would make inferential practice redundant) but also because it is unclear what sense there is in ascribing such a capacity. What range of skills could a speaker exercise to demonstrate grasp of these inferential connections? It would seem that all we can require of logical competence is an ability to recognize and perform correct inferences. But characterizing that ability will lead us to our second alternative because the most obvious way of doing this is to isolate a range of canonical inferential steps which would determine the use of the given sentence relative to a specific subset of other sentences.

Let us turn, then, to the second alternative. Understanding a sentence might be a matter of knowing its immediate (in a sense still awaiting full explication) logical consequences and/or grounds. But this position is hardly any more plausible than the last if it supposes that understanding a sentence would involve knowledge of every logically equivalent sentence, since what this means is that understanding a theorem would require understanding every theorem in logic. If that means that we need a knowledge of every theorem then we've regenerated the need for logical omniscience. If, though, it means that we should be capable of recognizing a theorem if presented with one or of checking whether a given construction is a proof or disproof of a putative theorem then, as we discovered in the last paragraph, we want some account of how we set about recognizing theorems. We can do that by syntactically characterizing canonical grounds and/or consequences. Canonical grounds and consequences thus become a combination of grounds and consequences given by the contents of the logically simple sentences and the syntactically defined rules of inference.

The position we now have is one where the meaning of a sentence is given by its canonical grounds and consequences where these are determined syntactically by the structure of the sentence itself. The canonical grounds and consequences will thus, in the context of logically complex sentences, determine when certain logical constants may be introduced and when eliminated. But this characterization of the use of the logical constants is not to be regarded as a characterization of their meaning. So what then would distinguish a semantics which takes the logical constants as implicitly defined primitives in the manner just indicated and a proof-theoretic semantics which takes the *meanings* of the logical constants as given by the introduction and/or elimination rules? Overtly it would seem very little. The difference would emerge in the spirit in which the canonical rules of inference were offered. In the second case we hold these to be responsible to antecedent semantical principles so we enjoin a requirement of harmony (however this is to be characterized) between introduction and elimination rules. The first position disclaims this reponsibility, rather, the rules of inference are simply explicit stipulations. In doing so, I shall argue, that it thereby exposes itself to the accusation that the explanation offered of a given sentence fails to confer on it a determinate sense.

Canonical grounds and consequences have, according to the first position, i.e., the implicit definitional view, thus far been defined purely as grounds and consequences which are determined by an explicit application of one of the introduction or elimination rules. Thus it would seem that we must mention *both* introduction and elimination rules in characterizing canonical grounds and also in characterzing canonical consequences. But then our explanations of, for instance canonical grounds, will inevitably be circular since the sentence will be assertible as a result of using the elimination rule on sentences of which it is a constituent. Thus the explanation will presuppose understanding of the sentence it is explaining. (e.g. P might be assertible as a result of using

MPP on $Q \rightarrow P$ so, in characterizing the meaning of P, we would need to give an account of this ground. But it is a ground for asserting P recognition of which presupposes an understanding of P since we have to characterize the assertibility of $Q \rightarrow P$ in terms of the assertibility of P itself.)

The alternative to this is to specify grounds in terms only of introduction rules. Thus the sentence becomes canonically assertible as a result of using the introduction rule of its dominant logical operator (so that in the example just given we do not need to refer to the assertibility of P as a result of MPP applied to $Q \rightarrow P$). We need, though, to find room for the elimination rules and this we can do by reference to canonical commitments arising from the elimination of the dominant logical operator. So canonical grounds and commitments correspond, respectively, to grounds given by introduction rules and consequences given by elimination rules. Canonical grounds and commitments would seem to match up but this is *not* a requirement on the adequacy of the stipulations and is precisely why both grounds and commitment need to be explicitly mentioned.

That they might not match up is shown by a consideration of the negation rules for classical logic. Consider a sentence of the form $\sim\sim P$ where P is a logically simple sentence of determinate meaning (given, say, by assertibility conditions). Then to understand $\sim\sim P$ involves knowing that it can be derived from P by double negation introduction (DNI) or by a reductio ad absurdum (RAA) with $\sim P$ as the sole assumption. But understanding also requires knowing how the logical operators can be eliminated according to our canonical rules of inference. Here we have double negation elimination (DNE) to give P. Thus P is assertible either when the original assertibility conditions hold or when we have a deduction of it from $\sim\sim P$. These conditions are distinct. This is important since an attempt to guarantee a match would be to attempt to justify deduction. Also, the nature of inference means that the deductive route provides conclusive grounds for *P*. So the deductive grounds cannot be regarded as grounds which merely provide good evidence that the sentence's previously given assertibility conditions obtain. Their function is entirely different from evidential grounds. But if this is the case then our previous characterization of the meaning of *P* was incomplete just because it failed to give an exhaustive specification of when it is correct to assert *P*. If we now attempt to include this ground in our account of the use conditions of *P* then once again our explanations become circular. An understanding of *P* presupposes an understanding of $\sim\sim P$ and that in turn presupposes an understanding of *P*.

One might wonder whether this limited holism is not acceptable since many concepts plausibly can only be understood in a context supplied by a group of similarly interdependent concepts; colour concepts might provide an example of this. But the present threat of holism is far more pervasive. The example was merely illustrative; nothing particular about negation was assumed. The reason the example was chosen was that intuitionistic arguments offer some *prima facie* reason for doubting whether the two sorts of grounds and consequences match up. However the same situation occurs whenever a logically simple sentence is assertible as a result of use of an elimination rule. (e.g., Consider the meaning of *P* as characterized by its assertibility conditions. Now ask whether assertion of *P* by use of &-elimination with *P&Q* as premise is included in that characterization. If it is then we develop a circularity since grasp of *P&Q* will involve grasp of the assertion conditions of *P*. If not, then it seems that we have an inadequate characterization of the meaning of *P*. The latter horn of the dilemma may be hard to take entirely seriously because of the obvious harmony between introduction and elimination rules. But this notion of harmony can have no role in the implicit definitional view.) If this is the case meanings become irredeemably indeterminate, Holism_3 collapses into Holism_1.

What if the situation is weakened to one in which we *assume* that a justification of the practice conforming to molecular requirements is given in an intuitionistic semantics, but that we introduce DNE by stipulation to recover classical logic? In other words we accept only a partial semantic justification of the logic and avoid the mismatch between what we are able to justify and the actual practice by introducing explicit stipulations. So, although some interest is invested in the justificatory programme, its revisionary power is sapped from the outset. Wright (1981, pp.19-21) focuses on just such an example and wonders why our original molecular view is prejudiced by the stipulation that two types of sentence with hitherto distinct assertibility conditions shall henceforth coincide,

> ... consider any language for which an assertibility-conditions semantics is correct whose effect is to open up, for certain statements, a possible hiatus between conditions warranting the assertion of their double negations and conditions warranting the assertion of those statments themselves. The question is: if a molecular view of such a language is possible at all, how would it be compromised if the language were merely altered so as to obliterate this hitherto recognised distinction? (1981, p.20-1)

The problem with this picture is that introduction of the stipulation means that we have no guarantee that we shall be able to maintain our grasp of the meanings as given intuitionistically. Here the threat is not of small local holisms but that the stipulation will alter the meanings of the intuitionistically given logical vocabulary resulting in wholesale changes of meaning in the fragment of language. Either we lose sight of the meanings as given intuitionistically -in which case the original problems recur- or else we retain the intuitionistic language in some form and use it to explain the meaning of the stipulation. The second scenario as well as being highly artificial does not promise a global reinstatement of classical logic nor does it forego a concern with semantic justifications. (Recall it is the question about the global applicability of classical logic that has been linked to the acceptability of realism.) It achieves at most a local rehabilitation of classical logic substantiated by some sort of reductive thesis. So we are not in the position of retaining a molecular view together with a refusal to account semantically for the logic; rather we accept the oblgation to account for logic, retain a molecular view and justify classical logic *locally*. That is a perfectly acceptable form of anti-realism about a region of discourse.

This point about the need to envisage a language given a prior intuitionistic semantics also shows that the appearance of a final position which is non-revisionary is entirely illusory. The language for which we accept the validity of classical logic has markedly less expressive power than the original language since the content of P in the original language has been weakened to that of $\neg\neg P$ (in the original language). There is no longer a means of expressing the content orignally conveyed by P.

§3 The Form of Dummett's Programme and the Importance of the Notion of Truth:

Wright's advocacy of what he terms the implicit definitional view of the meanings of the logical connectives is supposed to be consistent with Dummett's semantic constraints. Those constraints ushered in a general argument against a notion of truth possessing determinate truth values irrespective of our ability to determine those values. So Wright wants to show that we can accept the incoherence of realist truth but divorces the implications of that from possible revisions of logic. He attempts to achieve this by refusing to seek for a semantic validation of logic yet showing that that refusal does not invite the charge of holism. I have tried to show that Wright does not successfully distinguish his position from a holistic one. Now I want to consider a set of arguments drawn from suggestions of Wright which challenges the use Dummett makes of the notion of truth in characterizing the metaphysical issue of realism. These arguments pose a potentially deeper challenge to Dummett's programme since they question the semantic framework within which Dummett thinks the metaphysical issues must be tackled.

"The relation of truth to recognition of truth is," as Dummett notes in *The Justification of Deduction* (1975a, p.314), "the fundamental problem of the theory of meaning, or, what is the same thing, of metaphysics: for the question as to the nature of reality is also the question what is the appropriate notion of truth for the sentences of our language, or, again, how we represent reality by means of sentences." A discourse enables representation of an aspect of reality. To determine the manner of that representation we need to focus on the notion of truth relevant to that discourse. How do we arrive at the correct notion of truth? Dummett's answer is by attempting to construct a theory of meaning which is a theory of understanding for the discourse. The importance of this step is twofold. First, attending to questions about meaning allows us, without begging any metaphysical questions, to impose generally applicable constraints on an adequate account of our practice. This corresponds to the global aspect of anti-realism in that the anti-realist semantic arguments provide us with a generally applicable challenge to attempts to justify a practice. Secondly, the attempt to construct a theory of meaning for a particular region of discourse is constrained by having to give an adequate account of that practice. What counts as an adequate account is subject to debate. Certainly we cannot require that the account justify every aspect of the practice without some argument to the effect that such an account should be possible. The account must however engage with the practice itself in that we must show how semantic notions used in the theory relate to the practice.

Wright (1986) argues that accepting the principle of Bivalence for statements not known to be effectively decidable is not distinctive of a realist attitude towards those statements. He claims also that the anti-realist is not committed to a particular attitude towards the principle of Bivalence as applied to such statements. Before examining some of the detail of Wright's arguments I want, briefly, to investigate their import.

My immediate concern here is with the question of whether or not the search for a semantic justification of logic is a project which is both acceptable and necessary for an anti-realist. The ability of an anti-realist to accept the principle of Bivalence has no immediate bearing on that question. Before we can appraise the relevance of the anti-realist's attitude towards the principle of Bivalence we need to be given some account of how he maintains that attitude. However, even if we bracket that question for the moment there is, at least *prima facie*, a problem for the anti-realist. Realism involves a commitment to a notion of truth which determinately either obtains or fails to obtain independently of our ability to determine the matter. It might seem natural that accepting such a notion of truth should be explained in terms of an endorsement of Bivalence for all statements (and, in particular, for statements not known to be effectively decidable). Now, if it is questionable whether or not an anti-realist can endorse Bivalence, the nature of the metaphysical issue of realism becomes unclear. That is, the anti-realist faces a substantial task in merely *distinguishing* himself from the realist.

In some of his later work Dummett (1982, 1981, 1991) has retracted his over-simple characterization of realism with respect to a given class of statements purely in terms of the applicability of bivalence to those statements. The nature of the semantic theory adopted for those statements is itself an integral part of what constitutes a realistic attitude towards those statements. One is a realist about statements of a given class if one accepts a classical two-valued truth conditional semantics for those statements which makes essential use of the notion of reference (1982, pp.56-7). Or, better, such a position evinces an attitude of naive realism with respect to those statements. Dummett does not pretend that naive realism is in the least plausible as a global position. In many instances a realist may offer a reductive account of the truth conditions of one class of statements (the problematic class) in terms of those of another (the reductive class). Thus a truth conditional meaning theory is given for the former class which may legitimate the application of the principle of Bivalence. However a naively realistic position is not assumed since an inevitable appeal to the notion of reference for singular terms of the problematic class is not made in the semantic theory. It would, however, be incorrect to identify such a position with a relinquishment of a realist view of the problematic class. Dummett dubs such positions versions of sophisticated realism. (It is an important question whether sophisticated realism must always assume a naive realism with respect to some class of statements.) The metaphysical question now centres on reasons for departures from naive realism: no departure is metaphysically neutral. Anti-realism of Dummett's sort offers a globally applicable challenge to any semantic theory. Some positions, although not of the naively realist sort, will still fail to meet that challenge. An example of this might be certain many valued logical systems whose values determinately obtain or fail to obtain. *If* a naively realist view is at all plausible with respect to the class of statements concerned then such positions will be seen as departures from realism. (Thus will not be metaphysically neutral). But the reasons for departing from realism may be drawn from, say, a particular ontological view and thus the resulting position may fail to square with Dummett's constraints on an acceptable semantic theory. So Bivalence is only a necessary condition for a naively realistic attitude and a position may fail to be fully (i.e., naively) realistic yet still fail to meet the anti-realist's challenge.

We should note the following. First, it is not clear that an endorsement of classical logic will involve a commitment to Bivalence and secondly it is not clear that an anti-realist may not subscribe to (a statement of) Bivalence. So it may be that our first order linguistic practice may be preserved without evincing a realistic attitude. The relation between our constraints on a semantic account of the practice and the nature of the practice itself must be mediated by the complex business of constructing the semantic theory.

So Wright is correct in holding that there cannot be any *identification* of anti-realism with a specific attitude towards the principle of Bivalence. However that position goes no way towards showing that the anti-realist is unable to concern himself with semantic validations of logic. But, perhaps we can uncover a more profoundly unsettling problem for the anti-realist. Granted that the issue of realism does not hinge purely on the nature of the first order practice but essentially involves the nature of the semantic account of that practice, it may be a moot and significant point as to whether or not the semantic account is itself anti-realistically acceptable. In other words, if we are not sure whether or not use of Bivalence in the first order practice is legitimate, is it still clear whether or not the *semantic theory* makes use of an anti-realistically repugnant notion? If we cannot be certain of the credentials of the semantic theory the dispute threatens to become unapproachable.

So much then for the form of an attack on the coherence of (Dummettian) anti-realism. To take the question further I shall need to consider the specific arguments canvassed by Wright in his investigation of the role of Bivalence.

Wright argues that the anti-realist who adheres to intuitionistic logic can coherently either deny or accept Bivalence. So (semantic) anti-realism should not be regarded as *entailing* an agnosticism about

Bivalence. Obviously, if our concern is with distinguishing anti-realism from realism and Bivalence is taken to be necessary to realism, we need only concern ourselves with the possibility of an anti-realist endorsement of Bivalence. Wright notes that the anti-realist might reinterpret the statement of Bivalence as saying that all statments are (humanly) resoluble. He can thus accept Bivalence on faith without incurring a commitment to realism, that is, without incurring a commitment to the possibility of unknowable truths.

The target of these attacks is the view that an anti-realist is committed to an agnosticism about Bivalence. There is no obvious extension of the argument to motivate the sort of doubt about the coherence of the anti-realist position I sketched above. The scenario depicts the anti-realist as an adherent of intuitionistic logic and thus presupposes that a departure from classical logic has already been accomplished. That departure is presumably to be taken as a response to an anti-realistic validation of logical practice. Thus the legitimacy of that project must be granted at least as an assumption for the purposes of a reductio proof. But then it is clear that the principle of Bivalence is being accepted on faith and not as part of the anti-realist's semantic theory. The ability to make this distinction is all that is required to distinguish the anti-realist from the realist. Secondly the anti-realist interpretation of the principle of Bivalence (as evincing the belief that all statements are humanly resoluble, i.e., that all statements are anti-realistically true) must be regarded as intelligible. It can only be taken as intelligible if we regard the anti-realist's notion of truth (i.e., a concept of truth firmly tied to our recognitional abilities) as coherent. This, again allows the anti-realist to distinguish himself from the realist since the realist will not offer such an account of his bivalent conception of truth.

It might appear that a more powerful argument is to be gleaned from considering the relation of *classical* logic to the principle of Bivalence.

Wright (1986, pp. 351-2) argues that *if* classical logic is anti-realistically acceptable then it is not clear that espousal of the principle of Bivalence is not to be interpreted as an informal statement of the Law of Excluded Middle (LEM) -which, by assumption, is anti-realistically acceptable. One wants immediately to respond to this on behalf of the anti-realist that only for the realist is the principle of Bivalence a *semantic* principle used as a *justification* of the latter syntactic principle. But this distinction, Wright claims, is precisely what is required to be made out and, whilst it is clear enough for formal languages, requires explication before it can be applied to the case of natural languages.

The pertinence of this observation escapes me, Wright is claiming that in natural language the distinction between semantic and syntactic principles is not sufficiently well established to enable a separation of the use of the statement that all sentences are true or are not true as a statement of excluded middle from a statement of Bivalence. Perhaps so, but the *role* of the statement is, in either case different. In both cases the statement will be appealed to in order to justify *particular* inferences by excluded middle but only the realist will appeal to the principle of Bivience in order to justify LEM itself (which presumably, in its more orthodox form, is intelligible to both disputants). For the anti-realist such an appeal would be blatantly circular. Also the anti-realist is supposed to be in possession of an anti-realistically acceptable semantic validation of classical logic: it is to this that he will appeal in order to *justify*, rather than state, LEM.

Another way of putting the point would be in terms of speakers' knowledge of the relevant principles. Knowledge of the syntactic principle would be exhausted by an ability to apply the principle correctly. Knowledge of the semantic principle must be explained in terms of a grasp of the relevant semantical concept. This cannot, on pain of circularity, be taken to consist simply in the ability to perform and recognize appropriate inferences.

Wright stresses that the scenario described is premissed on the ability of the anti-realist to furnish a semantic justification of classical logic. That possibility might be thought to be implausibly far-fetched but both Wright and Dummett have provided super-valuational semantics which promise an anti-realistically acceptable justification of classical logic, at least in restricted contexts. The argument might be read as offering a *reductio ad absurdum* of a non-revisionary anti-realist stance. The argument would then be that if such a stance is assumed there can be no distinguishing the use *in the semantic theory* of anti-realistically repugnant notions. That result would be significant (partly because the apparent success of super-valuational approaches would then be *prima facie* evidence for the incoherence of the anti-realist's position).

An argument of this form is not, despite appearances, in the offing. Use of the statement that all sentences are true or not true as an informal statement of LEM is premissed not only on the validity of classical logic by anti-realist lights but also on the acceptability of the equivalence thesis. What we have, for any statement *P* is,

"*P*" is true or is not true.

i.e., "*P*" is true or "*P*" is not true.

i.e., "*P*" is true or not-("*P*" is true)

i.e., *P* or not-*P*, provided "*P*" is true iff *P*.

The argument I sketched above can now be seen to fail because in the super-valuational approaches the equivalence thesis is invalid because truth does not distribute over disjunction. The point is as follows. If the equivalence thesis holds for "*P* or *Q*" then we have,

"*P* or *Q*" is true iff *P* or *Q*.

But then, if the equivalence thesis holds for arbitrary *P* and *Q* (i.e., we have, "*P*" is true iff *P*) then we have,

"*P* or *Q*" is true iff "P" is true or "*Q*" is true.

So truth will distribute over disjunction contrary to the original claim.

Thus the most that the argument could show is that there is no anti-realistically acceptable validation of classical logic which also endorses the equivalence thesis (i.e., accepts the two valued truth tables as correct). But that, since it is a version of realism, should not be terribly surprisng.

It is conceivable that this argument might be objected to by insisting that we adhere to some non-standard interpretation of disjunction *in the semantic theory*. But, unless we assume that the connectives are given the standard truth functional interpretation (this is *not* to say that their *meanings* are captured in this manner) it is impossible to see how, indeed, we could characterize the notion of truth which the anti-realist finds unacceptable. It would, otherwise, be possible to continually bring into question the terms in which we characterize such a notion and we would thus never have a means of characterizing a notion which engages with the practice.

The possibility of the dispute would then threaten to vanish. But this only shows that the dispute requires that there be a measure of agreement between the disputants. We need only insist that the nature of the agreement begs no metaphysical questions. It is hard to see how insisting on the standard interpretation of the connectives in the semantic theory prejudges the metaphysical issue.

It should, finally, also be noted that the position the argument promises to deliver is that there is an incoherence in the anti-realist's position because he is unable to characterize those semantic notions which contravene his general considerations about meaning. This is not the position which Wright at some points seems to be advertising. Wright seems to be saying that we can accept the anti-realist's semantic constraints yet question the need for a potentially revisionary position to flow therefrom. The position rather questions the anti-realist's ability to formulate coherently those constraints. (We shall see below that Wright offers what purport to be more profound reasons for holding just this view.)

The point of my arguments thus far has been firstly to refute Wright's claim that abandoning the project of justifying logic is not to abandon an important conception of meaning, namely, that which insists on molecularity of meaning. Secondly I have tried to show that accepting the anti-realist framework does lead, contra Wright's contentions, to, at least, a potentially revisionary position. I have, however, implicitly assumed that one is able to substantiate a notion of meaning which allows that a certain pattern of use may or may not be faithful to or in harmony with established meanings. I have thus framed the argument within the paramenters of Dummett's anti-realist discussion. Wright contests whether these notions are themselves anti-realistically acceptable. This challenges the anti-realist's ability to state his own case and thus the very coherence of Dummett's position.

§4 The Possibility of an Anti-Realist Justification of Deduction:

If Wright is objecting to the anti-realist's use of notions to do with fidelity to pre-established meanings then the original portrayal which I gave of his position needs altering to the more forceful one of holding not simply that we can shift our methodological stance but that we must shift that stance and that that shift corresponds to an important realignment in our conception of meaning. The shift is not merely to a position which promises more, avoids the need for revision, at little or no cost. Rather we must shift because the original position is seen to be incoherent.

Wright objects to the anti-realist's ability to state coherently his claim that inferential rules are adequate only if they issue in warrants for the assertion of sentences in the base class that would pre-inferentially have been correctly assertible. The question Wright poses is whether we could ever recognize that a given practice of inference might fail this adequacy condition. If not then by his own arguments the anti-realist must admit that the adequacy condition proposed is vacuous. We might doubt our ability to recognize an inadequate inferential practice just because once accepted that practice serves a normative function. The rules themselves are involved in how we draw the line between appearance and reality.

An alternative arithmetic? To illustrate this let us consider an example which Wright discusses (1981, pp.21-7). Consider a community with exactly similar counting practices to our own save that they fail to go on to develop arithmetic. The community somehow finds itself in possession of a set of, what it takes to be, arithmetic truths and uses these as normative over its counting practice. Among the set of truths it accepts is the statement "17+29=45" (which we take to be false). We take the statement to be false because it will convict a correct count of, say, 17 then 29 and a total of 46 of error and will sanction an incorrect count of 17 then 29 then a total of 45.

The question Wright poses (1981, p.22) is how once the statement is accepted by the community as an arithmetic truth we could convince them that, in fact, it is not. If we cannot achieve this then our conviction about the correctness of certain counting operations will simply testify to the way we draw the appearance/reality distinction and not to the actual presence or absence of suitable pre-inferential warrants. On the basis of our arithmetic we say that although a count seemed correct there really was an error. But that same claim is available to the community in cases where their arithmetic calls on them to disagree with the result of a seemingly correct count. Because each claim is an open existential claim (that is, we do not give a precise specification of the set of possible errors) each is strictly indefeasible. So in no case of divergence can we convince them (or vice-versa) that their arithmetic is at fault. Thus the adequacy condition on the inferential practice is empty (or, at least the only way to give substance to it is to say that despite our inability to find an objective fact determining the correct arithmetic there nevertheless is some such fact -and that is anti-realistically unacceptable). So the anti-realist constraint of conservativeness is vacuous.

The conclusion that anti-realist use of the notion of conservativeness is illicit does not however immediately follow from this illustration which reveals merely that *in any given case* we are incapable of determining whether the rule applies conservatively or not. An overall constraint on an appearance/reality distinction is that most of the time what seems to be the case is the case. If we accept this principle then observations of many counts of say 17 then 29 then the total will result either in us saying that what seemed to be the case mostly was not the case or else they must say this (or else there might be a "tie" in which case we had better give up doing arithmetic altogether). So the non-conservativeness of the rule will emerge in its being a bad arbiter of reality. Two flaws in this view seem to be that: it makes arithmetic seem empirical; and, the original principle lacks motivation.

To take the second objection first, let us assume (as I feel confident of this being the upshot of the "experiment") that it is the community rather than us who get things wrong most of the time. The question is; what explanation can they offer themselves of this situation? If they feel no explanation is called for then we would surely take them to be doing something approaching a ritual of no intrinsic significance to them. Arithmetic occupies the place it does in our lives partly through its ability to generate precise expectations of which we are thoroughly convinced. Otherwise we would not be taking it as an arbiter of what really is the case. If no explanation in the face of massive failure is called for then the inferential practice must occupy a very different rôle in their lives to that in ours. So we would not be comparing similar practices. Admission on their part that their arithmetic is at fault is just the result we're after so short of that answer what explanation can they give themselves?

Given an instance of miscounting three sources of error suggest themselves: a lapse in cognitive faculties; a misconception in the nature of the objects counted (we're mistaken in believing they persist through time or that they're distinct etc.); or, a misunderstanding of counting procedures. Globalised in the way they need to be, none of these explanations is in the least plausible. The hypothesis of a global lapse in cognitive faculties is just to say that we are cognitively limited. So that in the case in question we cannot implement the counting procedure. Now, unless we are to allow this position to inflate into a blanket scepticism about whether we ever can have knowledge of having satisfied operational criteria there must be a range of cases where such knowledge is (defeasibly) available. In these cases the global hypothesis of a cognitive lapse is unavailable. The position that emerges from this consideration is a strict finitistic one, i.e., we accept that we need to justify inferential criteria only relative to our actual ability to implement operational criteria. I am not endorsing that position. What I am saying is that the possibility of such a position shows that grasp of operational criteria allows there to be some question of whether arithmetic criteria are acceptable. In particular, it is possible to notice a clash between the two sets of criteria. However, having thus demonstrated the possibility of justification, I shall go on to argue that the programme results not in strict finitism but in intuitionism.

Globalising the third alternative very quickly leads to absurdity since it is just to suppose that we misunderstand our own language.

What of the second alternative? This amounts to saying that our use of observational criteria for counting, which hitherto had seemed to be applicable, are shown by the inferential criteria not to be so in reality. So conditions which we had taken to warrant the assertion of numerical predicates no longer do so. Since these conditions determined the meaning of these predicates those meanings must have altered. Thus the community will, in this case, find that the inferential criteria extend the observational criteria non-conservatively. So the community cannot satisfy itself that where inferential criteria result in massive accusations of error those criteria are, nevertheless, adequate. I conclude, therefore, that we must accept as a constraint on an acceptable means of distinguishing appearance from reality that we are not forced by the distinction to accuse ourselves of error too often.

Returning now to the first objection to my adequacy condition on an appearance/reality distinction: does the description given make mathematics seem empirical? Certainly it shows that the applicability of mathematics depends on empirical regularities. But that is different from saying that we accept mathematical statements *because* they coincide with observed regularities. The manner in which we come to accept mathematical statements, i.e., via proof, is crucial. In the example we supposed the community not to have proofs for the arithmetic statements it accepted. There are two difficulties with changing the example so that the community possesses what purport to be proofs. The first is that the nature of counter-arithmetic proofs is utterly mysterious to us. Secondly if they did possess such "proofs" then we could discuss these and discover precise disagreements traceable to differences in meaning or in our account of meaning. That is, either we will agree that they have a divergent but meaningful practice (which we might choose to adopt, having learnt it) or else we shall disagree about how we imbue our terms with meaning.

"Laying down a measure" (to use Wittgenstein's phrase) may depend upon definitely expecting certain outcomes. Disappointment in those expectations can lead to us not using the measure. To this extent a mathematical theory is empirical: but for the sure expectation (which is a contingency) it would not be a measure. But just because the mathematical proposition is used as a measure, tells us what we can expect with certainty, it is not empirical. Thus the mathematical proposition does not itself express the sure expectation although adoption of it depends upon there being sufficient regularity in results which lead to some certainties. (Wittgenstein, 1978, IV §§ 52,3)

The argument shows that a condition of adequacy on inferential criteria is that for statements decidable by operational criteria we are not forced to accuse ourselves of error too often. This condition is weak. It, in effect, amounts to saying little more than that no-one would adopt a duff arithmetic. Wright cannot be taken to be suggesting that they would. What this banal consideration, if that it be, does achieve is a motivation of our, as it were, realist intuition that there can be at most one acceptable arithmetic. The thought experiment must be accepted by anyone engaging in inference so no discord in inferential practice can be accepted with equanimity. Moreover our reason for insisting on agreement in inferential practice does not depend on the thought that one practice is bound to have a practical superiority over the other. It is not that in the process of use we discover one practice to be better than the other (it may be that we never make enough use of the divergent results for this to emerge) but that the thought experiment gives us *a priori* reason for refusing to accept the correctness of both systems or, in other words, for affirming that, at least, one system must be absolutely *in*correct. The thought experiment shows an incoherence in conceiving that more than one arithmetic might be correct or that judgements of correctness or incorrectness are always made from the perspective of accepting one or other inferential practice so that, strictly, they are incomparable.

§5 Anti-realism and Soundness:

Wright shows that any (consistent) inferential practice is bound to seem sound to those engaged in it. Combine this with the view that at most one inferential practice is, in fact, sound and we land ourselves in the predicament of holding that there is a substantial question about whether or not our inferential practice is actually sound and no means of satisfying ourselves one way or the other.

Wright argues that since, in any given case, a disparity in the verdicts obtained by use of operational criteria and inferential criteria can be explained in terms of there having been a mistake in the implementation of the former there can be no question of our operational practice "forcing" a given inferential practice on us through a general constraint of soundness. Or, better put the argument runs as follows. The anti-realist wants to concern himself with the question of soundness of a given inferential practice. Now, in any given implementation of inferential criteria we cannot find ourselves in the position of *having* to see that implementation as unsound since the *appearance* of unsoundness is all we ever arrive at and that appearance is always explicable in terms of the strictly indefeasible hypothesis that there was some mistake. Thus the unsoundness of the system (if it be so) must be a feature of the system which evades recognition. So, in claiming the right to be concerned with issues of soundness one is committed to there being a "determinate objective class of cases" in which a mistake either did or did not occur, independently of our ability to determine the matter. This is a position of doubtful coherence for an anti-realist to adopt given his own arguments against the possibility of our grasping transcendent states of affairs which determinately either do or do not obtain.

The argument needs something of a gloss. It shows that the unsoundness of a given inferential system is unrecognizable. Or, rather, that it only can appear to be unsound from the perspective of an opposing system. But, since the argument is symmetrical, there is no reassurance to be had that that system is actually correct and so no absolute accusation of unsoundness. This, given a loose interpretation of the intuitionistic position, might be held to be tantamount to holding that the system is not unsound. So *any* system has been shown by intuitionistic lights to be not unsound (contradiction aside). Wright promises us an argument that the anti-realist cannot coherently concern himself with questions of soundness. That argument presumably emerges in the following manner. Assume that A_1 and A_2 are conflicting inferential systems in the sense that they issue in contradictory propositions. So if A_1 is sound A_2 is unsound and conversely. We have just shown that, by intuitionistic lights, both systems are not unsound. But then, were we to have a proof of the soundness of A_1, we would have a demonstration of the unsoundness of A_2, which would contradict the fact that A_2 is not unsound. So we can have no proof of the soundness of A_1. This, provided we allow the anti-realist to be concerned with questions of soundness, again leads to contradiction since then we have, intuitionistically, that A_1 is unsound. So we conclude that the anti-realist can have no legitimate concern with the notion of soundness.

Note, first, that the argument relies on a loose interpretation of the intuitionistic logical constants since, strictly, to conclude that any system is not unsound we should have to derive a *contradiction* from the assumption that it can be recognized as unsound. We have only been given a demonstration of a certain incoherence attaching to this supposition. Secondly, the argument requires the above gloss since, in its absence, all we have is a general argument to the effect that we have no effective means of recognizing the unsoundness of a given system. But the lack of a negative decision procedure is not enough to preclude legitimate anti-realist concerns with questions of *soundness*. That is, the argument falls short of forcing on us the conclusion that we are unable to demonstrate that a certain system is sound. The latter might still be achievable despite our not being able to show that a given, consistent system is unsound.

An important step in the argument as I have reconstructed it relies on an assumption about the nature of the clash envisaged. We are asked to imagine alternative systems where the soundness of the one entails the unsoundness of the other (since we are supposed to be able to prove in one the contradictory of a proposition provable in the other). We need not however only be concerned with clashes of this form. The issue which concerns us most, that concerning the soundness of classical and intuitionistic logic, does not take this form. The intuitionist does not accuse the classical logician of adhering to an unsound system of inference, rather, he questions the scope of certain rules of inference. That is, acceptance of intuitionistic logic does not entail that one see the classical system as being unsound (relative to the *existence* of non-inferential warrants), indeed the intuitionist might agree that classical logic is not unsound, but (in view of his adherence to intuitionistic principles) can still ask for an assurance that classical logic is indeed sound.

But, before investigating that dialectic, I should clarify the present position. It would seem that raising the question of the legitimacy of an underlying logic is premature since, as the argument stands, the anti-realist cannot concern himself with questions of soundness of conflicting inferential systems. And, if he cannot motivate a coherent interest in soundness in cases of that sort, how can there be for him any substance to the notion of soundness which requires fidelity to established meanings or non-inferential practices of assertion?

The example we are presented with works by asking us whether or not we could recognize the unsoundness of a system which issues in a counter-arithmetic result. A premiss is that we agree with the deviant arithmeticians in our practice of counting so that there can be no suggestion that the discrepancy is explicable as arising out of divergences in the meanings of our respective numerical terms. In a sense we have a single community of counters who diverge amongst themselves as to which arithmetic truths they should endorse. Or, put another way, the community has a contradictory arithmetic. Admittedly, no single speaker will find himself (given the terms of the example) wanting to endorse both of two contradictory propositions. But this is because we imagine a certain asymmetry: advocates of the truth of 29+17=46 are thought of as possessing a proof of that result whereas advocates of the truth of 29+17=45 are not. It is supposed that anyone with sufficient nous to follow a proof of the former will see his rashness in accepting the latter. But if this supposition is held universally true then it supplies a reason for holding our arithmetic to be correct: the disagreement seems trivial (this position is consistent with holding that there are people in the community who are simply too dim to grasp the proof and so continue in their deviant ways). If it is not held to be universally true then we are asked to imagine speakers who grasp the proof yet do not relinquish their belief in the former (false) proposition. But in this case their practice is clearly contradictory. Thus the grounds for holding that the deviant arithmetic is unsound stem from the fact that it is inconsistent. Neglect of this aspect will, unsurprisingly, render us incapable of recognizing the system to be unsound.

The example is itself deceptive in that it calls on us to imagine a foreign community with a conflicting practice. When phrased in this way it seems that their practice is consistent and alternative. We then seem to have a separable task (which according to the argument we cannot perform) of showing that the system is unsound. But this is misleading. We should rather view the position as one infecting a single community which comes to see its practice as inconsistent and thus in need of reform. Of course, the precise nature of the reform may be a delicate matter, but the recognition of unsoundness was all we were asked to demonstrate and this I urge is possible through the recognition of an inconsistency.

It is apt to seem that in invoking the question of a linguistic community in this way I have simply and fairly blatantly gerrymandered the example to suit my ends. Naturally, I do not think that this is so. Let me briefly explain the underlying thought. If we are to imagine a genuinely foreign community then we must make sense of an interpretative exercise. We cannot simply assume that our understanding of numerical terms agrees with theirs without unpacking that assumption. If the assumption is based on a coincidence of counting practices then the example asks us to imagine a single community where counting practice leads to a contradictory inferential practice. If not, then we must interpret their counting practice. We would then want some explanation of that practice (this is especially true if the community held absurd arithmetical beliefs such as 29+17=45). That is, we would expect them to offer or to recognize something like the Peano axioms as explanations of their practice. Once they did this our grounds for rejecting their practice would be that it is demonstrably inconsistent. This may well raise further questions about methods of demonstration but such questions of the soundness of the underlying logic are just what concern us. My contention is that such questions can be legitimately raised by the anti-realist provided the soundness of the one system does not imply the unsoundness of the other. The point seems to be that the possibility of raising the dispute requires a measure of agreement, not that the dispute cannot be raised.

If this argument holds then we have no argument for the view that the anti-realist cannot concern himself with questions of soundness. So it is open to the anti-realist to question whether classical logic can be shown to be sound (even though he admits that it is not unsound). The point is that classical logic can be shown to be sound relative to truth as realistically conceived, and cannot be shown to be unsound relative to the mere existence of non-inferential warrants. (The latter follows from the former and from the specifically anti-realist considerations advanced by Wright.) The question though is whether we are justified in ascribing grasp of realist truth to ourselves. The anti-realist gives a general argument drawn from considerations about meaning to show that we cannot be justified in such a claim. In contrast to the realist, the anti-realist offers an account of an epistemically constrained notion of truth and asks whether classical logic is sound relative to that notion. But that question only asks us to reflect on our epistemic position at the conclusion of certain classically valid inferences. If we classically infer a proposition by using double negation elimination on a verification transcendent proposition or by using an argument by dilemma on the same then the intuitionist asks, "Does the possession of an inferential warrant for that proposition guarantee the obtainability of a non-inferential warrant?", i.e., in such cases do we have a method for obtaining the appropriate non-inferential warrant? The answer the intuitionist urges in response to that question is simply "no". This failure is clearly recognizable. The legitimacy of the anti-realist's question requires only that we be able to reflect on our epistemic position. Given a classical inferential warrant we do not need to settle its validity by determining whether or not a non-inferential warrant *really* exists, we simply need to question whether or not we are in a position to obtain such a warrant.

The programme precisely avoids the question which invites Wright's attack. That is, we do not put ourselves in the position of asking whether or not inferential criteria *actually* coincide with non-inferential criteria (and thus close a circle of dependence, since our verdict on the actual satisfaction of the latter is normatively constrained by inferential criteria themselves). We simply ask whether or not, when we possess inferential warrants for the assertion of a proposition, we have a means of obtaining non-inferential warrants.

§6 The Objectivity of Meaning:

In the last section I argued that considerations drawn from the normative role of inferential criteria do not suffice to preclude legitimate anti-realist concerns with questions of soundness of a given inferential practice; those arguments suffice to show, at most, that questions of *un*soundness of an ostensibly consistent system cannot be raised. I take this conclusion to indicate a deeper point of divergence between the Dummettian anti-realist and his Wrightian cousin. Wright's interpretation of the Rule Following Considerations of Wittgenstein lies at the heart of the matter and explains Wright's conviction that the notion of soundness is unintelligible.

Let us return to our arithmetic example. No conflict between arithmetic criteria and operational criteria, it was shown, could come to light since the arithmetic criteria are used to determine instances of misapplication of the operational criteria and the hypothesis of an error in such an application is strictly indefeasible. Wright claims that it makes no sense to question whether in such a case there *really* was a misapplication of the operational criteria which could have been recognized by pre-arithmetic criteria alone. So it makes no sense to question whether the arithmetic criteria are *really* sound with respect to counting procedures. Thus it is not simply that the question of unsoundness is undecidable but that the notion of soundness to which we appeal here is unintelligible.

The apparent need to ask questions of this form stems from a misrepresentation of the character of our own understanding. We assume that, in giving meaning to our terms, we are able to confer on them a determinate pattern of use such that conformity to this pattern of use would have required or would require a definite response in circumstances where we have not yet used the expression or cannot, in practice, so use it. Thus a concern with soundness betrays a belief that our understanding of counting procedures enjoins a correct pattern of implemention of them independent of our, actual or possible, use of those procedures. So that when we use an arithmetic criterion to determine an error it is perfectly determinate whether or not there is actually an error by criteria governing counting operations. Instead Wright claims all we can do is *say* that there was such an error but since, irrespective of the arithmetic adopted, it is always possible to resort to this saving hypothesis this is not to enter into any substantial commitment.

I cannot offer a response to Wright's treatment of the Rule Following Considerations which his thorough presentation deserves. What I want to do here is to outline my reasons for thinking that Wright's application of rule following is at fault here and thus that the anti-realist does not involve himself with a notion of soundness elucidated as above.

I am in complete agreement with Wright's eminently lucid interpretation of the Rule Following Considerations as providing a fundamental critique of attempts to account for our use of language in terms which "platonize" meaning, that is, which make our use accountable (in a normative sense) to some item (a rule. or a relation of reference or truth etc.) which we come to know. If manifestation is a general constraint on ascriptions of knowledge then the Rule Following Considerations show that knowledge of the required sort cannot be publicly manifested. So, if this knowledge is to provide our semantic basis we must make sense of the idea of private self-ascription of knowledge. The question then is whether each of us can manifest this knowledge to himself. The private language argument then shows the impossibility of the private linguist satisfying himself on this point since he cannot make a distinction between circumstances where his use appears to coincide with his putative knowledge and circumstances where it actually does so. So no self-ascription of knowledge of the semantic primitive is possible. In particular, if we think of language use as explained through grasp of truth conditions where truth is taken as primitive we can give no satisfactory account of our knowledge of these conditions.

Our search for a theory of meaning now takes the following turn.

First, we recognize that knowledge of any property chosen as central in the semantic theory must, necessarily, be *publicly* manifestable. Secondly, the Rule Following Considerations show that this knowledge cannot be the ultimate ground of correct use. A certain base class of expressions are given explanations which situate the use of the expression within a normatively constrained practice. The position now allows for Dummettian concerns about the nature of truth. We must give a reductive account of truth, the question is whether we can reinstate a notion of truth which globally applies bivalently. Dummett shows truth conceived in this way is of the essence of realism and that such a notion of truth is bound to contravene the manifestation requirement. (A reductive account based on compositionality would fail to reinstate classical truth because of, as Tennant (1987, p.112) observes, Gödelian reasons relating to incompleteness: any circumscription of the range of capacities we manifest relative to our grasp of a certain class of statements formed from a limited vocabulary will always fail to warrant grasp of a decision procedure which is, theoretically, unrestrictedly applicable.) Taking assertion as central in the semantic theory enables us to characterize our use of the truth predicate and so to reveal our notion of truth.

For Wright we must, as a result of the Rule Following Considerations, talk of meaning in terms, not simply of use, but as use against a "securable background of communal assent". Only in the presence of this background can we guarantee that our use is subject to normative constraints. So we can do no more than point suggestively to an aspect of the practice and say "This is what we do." The task of the philosophy of language is then to weed out confusions which result from theoretical "overloading" of the notion of meaning. An example would be the set of criticisms Wright levels at the Dummettian anti-realist. Notice that "first order" use of language, e.g., our claim that in the event of a counter arithmetic counting result we must have misapplied the counting procedure, is exempt from criticism.

There is a tension in Wright's position since, given an acceptance of his description of meaningful use, it is possible uncontroversially to talk about the assertion conditions as giving the meaning of sentences provided we do not use assertion conditions as objects of knowledge which *always guide* our use. It is not clear what Wright takes an assertion condition to be; the two options available to him correspond to the horns of a dilemma. Either a sentence is assertible just when we can secure the necessary communal assent, so that the assertion condition corresponds precisely to the condition of having communal assent, or assertion conditions are conditions in the world, that is they are conditions which the world fulfils and a recognition of which is taken as an entitlement to make the appropriate assertion. The first interpretation admits that we can command no view of the specific nature of meanings in our language since, given any systematic account of our use of language we shall have no grounds to distinguish it as a normative description rather than description of a regularity. This is simply because it will only systematise what the community has done and, since Rule Following Considerations apply to the community as a whole, this can have no normative influence on what it goes on to do. Whilst, on the second interpretation, if we say that a counter arithmetic result betrays an unnoticed error in applying counting procedures then there is no opportunity of dismissing this as just what we say. If we are entitled to say this then there must *really* have been a mistake. We do not have a stronger standard of objectivity to which the reality of the mistake is accountable.

But then what of Wright's demonstration of the relativity of such a judgement? The line I tried to urge a few paragraphs ago was that the relativity is a consequence of allowing the validity of justificationless judgements based on inference. We emerge from the tension if we allow that there is a base class of logically simple sentences for which

Wright's description of understanding is inevitable. That is, there is a class of expressions which can only be explained by demonstrating our use of them in the world. Rule Following Considerations are not intended to be revisionary about contents (non-meaning-theoretic contents, at least) which we take ourselves to grasp nor do they reduce to absurdity the idea that we succeed in normatively constraining use. The impact of the Rule Following Considerations lies in the criticism they offer of certain explanations of how correct use is determined. Now with respect to some base class of expressions, we can take it that we simply have an understanding of those expressions and then, given that understanding, attempt to characterize our notion of truth by looking at our justification of inferential practice when inserted into this base class. The problem with this (modest) approach is that we need to give some account of what the understanding of these sentences consists in in order to determine the central notion of the semantic theory which inference has to preserve. But the description of our grasp of these basic expressions can be given without presupposing any conceptual expertise since the use is grounded in a non-linguistic practice which is itself subject to normative constraints. (Recall the discussion of modesty and full-bloodedness in the last chapter.) The theory of meaning does not itself attempt an explication of the origin of the standards of correctness governing this practice, rather it simply describes them. So Rule Following Considerations are not applicable to the account we give of our grasp of these expressions just because that description presupposes, and fails to explain, standards of correctness.

To put the point differently, Rule Following Considerations show that understanding cannot always be described in terms of an internalisation of a rule or interpretation. So, if our account grounds the use of certain expressions in actions performed within a practice or custom, then the requisite notion of grasping a pattern of use will not present itself for an attack based on Rule Following Considerations. The practice or custom and its normative requirements are not given a norm-free description. This is not in order simply to dodge the Rule Following Considerations, rather it accepts as a *consequence* of the Rule Following Considerations that no complete explanation is possible.

Given this account we then insist that introduction of an inferential practice be justified only if it is sound relative to a semantic notion which can be taken to describe our understanding of the base class of expressions. The project thus consists of: first, accepting the nature of the practice as given by the contents expressible in the base class of expressions; and, secondly, questioning whether manifestation of grasp of a given content justifies the belief that it has a determinate truth value.

Wright pointed to a relativity in judgements of error and so criticised our conviction that it is objectively true that a counter inferential result could have been recognized pre-inferentially as in error. This relativity vanishes since either there is a divergence in judgement about the correctness of a use of (logical or of non-logical) expressions and one or other disputant lacks a justification or both possess putative justifications. If the divergence relates to justificationless use of expressions in the base class then there is a disagreement about the meaning of those terms, i.e., we are considering different, not *alternative*, discursive practices. If the divergence relates to an aspect of their logical practice then, since we are allowing play with the notion of having a justification, lack of justification is to be impugned. Lastly, the divergence in justifications of logical expressions can only relate to the manner of representing the content of expressions in the base class to ourselves. But this regenerates our concern with a semantic justification of logic based on a suitable central notion: we resurrect the semantic debate.

§7 Summary:

Wright is arguing for two conclusions: the first is conditional; the second is not. He wants to show, first, that, given his own arguments, the anti-realist cannot motivate an interest in semantic justifications of logic and, secondly, that the Rule Following Considerations show that there can be no legitimate concern with semantic justifications of logic. I questioned the first claim simply by showing that Wright's argument depends on importing his own interpretation of the Rule Following Considerations and that there is an alternative interpretation consistent with anti-realism as I see it. Wright takes the anti-realist concern with public manifestability to be equivalent to his notion that meaningful use must always be described as use against "a securable background of communal assent" (since achieving this consent *is* to be correct). I think that this is wrong: the anti-realist needs to give some account of the pre-conditions for having a practice which involves norms but need not give any reductive account of what that correctness consists in. So there is a coherent anti-realist perspective which allows the metaphysical issue of realism to be raised in an orthodox Dummettian manner. To make a judgement on the second of Wright's claims I would need to settle the issue of how the Rule Following Considerations *must* be interpreted. But that issue takes me too far from present preoccupations.

I began the chapter by questioning Wright's claim that the requirements of molecularity do not demand that a semantic justification of logic be given. I gave reasons for thinking that any position refusing the project of giving a justification of logic will collapse into holism. I went on to defend the programme of giving a justification from attacks which questioned the anti-realist's ability to give substance to the semantical notions he intends to use (and which thus question his ability to distinguish himself from the realist). My point here was that the anti-realist can insist on sufficient agreement between the disputing parties to ensure that his semantical notions engage with first order practice without begging any metaphysical questions.

So my conclusion is that the anti-realist has a legitimate concern with semantic validations of logic and that the requirements of molecularity necessitate such a concern. Classical logic is thus only acceptable if it can be given an anti-realistically acceptable validation.

CHAPTER FOUR: STRICT FINITISM

Strict Finitism

I shall be considering the relation of three positions which can each be summarized as follows,

Strict finitism: Truth values of sentences are determinate when they can, in practice, be determined.

Intuitionism: Truth values of sentences are determinate when they can, in principle, be determined.

Realism: Truth values of sentences are determinate (irrespective of our ability to determine them even in principle).

In this chapter I want to defend the intuitionistic position from strict finitistic attacks which claim that the intuitionistic position, in its use of notions such as verifiable *in principle,* is susceptible to criticisms which are precisely analogous in form to those brought by the anti-realist against the realist. Accordingly, anti-realistically inspired intuitionism thus incoherently demands a revision in classical mathematics which is less extreme than that demanded by its own motivating lights.

§1 The Analogy Between the Intuitionistic Attack on Realism and the Strict Finitist Attack on Intuitionism:

Let us consider then why the intuitionist's notion of truth is thought to suffer from the same incoherence as that of the realist. The realist believes we confer a meaning on the terms of our language in such a way that any expression fit for making an assertoric statement is guaranteed a determinate truth value whether or not we have any method for determining that value. The anti-realist challenges this belief by asking what in the nature of our understanding of logic-free language justifies us in assuming that we grasp this bivalent conception of truth and by questioning how practice with certain rules of inference could possibly *constitute* grasp of this notion of truth. The point about the first challenge is that nothing in the conditions in which we learn to use and in which we correctly use an expression can justify assuming that conditions which lie so radically outwith our recognitional capacities determinately either obtain or fail to obtain. Knowing how to use an expression correctly is to know its meaning so nothing in the understanding of an expression can relate to the determinacy of such recognition transcendent states of affairs. The point is not that we fail to form the conception of definite truth conditions which lie beyond what we can guarantee to be able to recognize but that where we do form such a conception we cannot justify the assumption that such conditions determinately either do or do not obtain. That is, the challenge is directed at the assumption of determinate truth *values* not at the conception of truth conditions which outrun what we can (presently) recognize to be true. Insofar as we characterize the meaning of a sentence by its assertion conditions we give an acount of when it can justifiably be held to be true. So we give an account of its truth conditions. This applies equally to both effectively and non-effectively decidable sentences. So that in the latter case we can say that we know what it would be for the sentence to be true although we cannot guarantee either to put ourselves in a position of recognizing it as true or to show that it is impossible to do so. So, in this sense, we grasp truth conditions which are recognition transcendent although, since our notion of truth is epistemically constrained, i.e., the truth of a sentence will always be linked to ways in which we recognize it to be true, truth conditions are not radically recognition transcendent. Thus the notion of unknowable truths is repudiated. The effect of rejecting this conception of radically recognition transcendent truth conditions is, though, to reject the determinacy of truth values for non-effectively decidable sentences (and thus to reject bivalence as applied to them). One might say that we do not question which truth conditions we grasp but whether the nature of that understanding justifies holding that the truth condition is determinately fulfilled or not.

This point has been well made by McDowell in *Truth Conditions Bivalence and Verificationism* (1976) where he argues that one can accept the verificationist argument simply by accepting an intuitionistic proof theory in a Tarskian truth definition so withholding assent from the principle of bivalence but using a verification transcendent notion of truth. It is true that McDowell takes this as an argument for realism where he identifies that position with maintaining that truth may transcend verification. But realism is in fact a position relating wholly to the question of determinacy of truth value (see for instance, Tennant 1987 Ch. 11 and Luntley 1988 p.3ff). The point, as I have just remarked, is not whether or not we grasp truth conditions which (presently) transcend verificatory procedures but, when we do so, whether manifestation of that grasp justifies a belief in determinate truth values.

What then is the parallel argument against the intuitionistic belief that truth value is determinate whenever we have an effective decision procedure? Presumably this, that we can only justify our notion of truth relative to use that we are *actually* capable of making so we can only assume that truth values are determinate where we can actually determine those values. Nothing in the nature of this use shows that we grasp a notion of truth which is determinate in cases where we are not actually capable of instituting the decision procedure. So the intuitionist makes precisely the same assumption as the realist, namely, that determinacy of truth value outruns our ability to determine truth value. In the realist case this is extended beyond all possibility of determining truth value whereas in the intuitionist's case it is only extended to some possibility of determining truth value in principle. In either case, the strict finitist holds, the extension is unjustified by our grasp of the use of an expression.

A number of commentators (e.g. Wright, 1982; George, 1988) attempt to bring out this parallel by asking how the intuitionist can communicate his notion of truth, i.e., verifiability in principle, to the strict finitist. Just as the realist in his attempt to explain his notion of truth to the intuitionist is frustrated by the circularity of his attempted explanations so the intuitionist is frustrated by circularity in *his* explanatory programme. It should be noted that where we talk about communicating a notion of truth we could (should) be talking about giving a non-circular justification of the notion of truth. Take the realist's justification for assuming that quantification over an infinite domain (of a decidable predicate) always issues in a sentence with determinate truth value. Each instance of the quantified sentence, the realist will uncontroversially (relative to the intuitionist) maintain, has a determinate truth value so the infinitary product or sum of these truth values will itself be determinate. But in this last step the realist assumes just what he set out to show, namely, that truth values are determinate irrespective of our abilities to determine them.

How can the intuitionist set about explaining his notion of truth? The intuitionist wants to justify holding that truth values are determinate when determinable in principle. Say he puts forward an argument similar to the following; we have a procedure which is guaranteed, after a finite number of steps, to come to a halt issuing in a verdict on truth value. So whether or not we implement or can implement the procedure it must be determinate in advance of instituting it which truth value it will, if correctly carried out, issue in. Thus the truth value of the sentence is determinate. The strict finitist can attack this explanation at two points. First, he might agree that if the procedure is surveyably implemented it issues in a determinate verdict on truth value but question whether we are entitled to assume that it is determinate which verdict implementation of the procedure will issue in. Only in the latter case can we take it that the truth value is determinate but to move from the former to the latter is just to assume that in principle determinable truth values are determinate. Secondly, the strict finitist can question the intuitionist's notion of finite number. If the notion of finitude coincides with something the strict finitist is prepared to endorse then it might be that a finite procedure as used in the explanation of "in principle decidability" coincides with actually implementable procedures. In this case the intuitionist position collapses into the strict finitist position because he is simply unable to convey the peculiarly intuitionistic notion of "in principle decidability" to the strict finitist.

We can again look at the analogy between the realist and intuitionist by noticing that both insist that they have a procedure for determining truth value. The realist's method consists in surveying an infinite domain of truth values while the intuitionist's method consists in implementing a finitely bounded process. Ignoring for the moment questions about what the intuitionist means by finite, it still appears that both descriptions are similarly question begging. We want the realist to explain why an undecidable sentence, i.e., the result of surveying an infinite domain, admits of having a determinate truth value and the intuitionist to explain why the outcome of a finite procedure should be determinate. What, one might ask the intuitionist, is the relevant difference between the infinite and finite but unsurveyable case? After all surely the *essential* problem with the infinite case is just that infinite domains are, in practice, unsurveyable.

§2 Strains in the Analogy:

In this section I want to question how deep the analogy between the intuitionist attack on realism and the strict finitist attack on intuitionism runs. My thought is that by focussing on the extent of agreement between disputing parties we can reveal a crucial disananolgy between the two cases. Of course, merely showing that there is a disanalogy between the two cases will not constitute a defence of intuitionism. A full defence would have to show how the disanalogy undercuts the strict finitist attacks on intuitionism whilst leaving untouched the intuitionist's attack on realism. My argument is considerably weaker and less direct than this. Having demonstrated a disanalogy I then go on to argue that the strict finitist position is simply not viable and must be rejected: we cannot accept the strict finitist's notion of truth if we are to allow for the practice of inference. It might be thought that one taking this attitude would be accepting a *reductio ad absurdum* of anti-realism viz., that anti-realist arguments must be wrong since full implementation of those arguments leads to adoption of an incoherent position. The point I want to make, however, is that importing the *general* constraint that our account must allow for the practice of inference, demonstrates a need to accept a notion of truth which determinately obtains even when we cannot, in practice, determine its value (in Dummett's words we are forced by the practice of inference to make some "concession to realism") but that, because of the disanalogy I claim to have demonstrated and the evident weaker assumption involved in intuitionism no direct argument for the reintroduction of full-blown realism is forthcoming. Intuitionism is thus (weakly) vindicated as the minimal viable position. That is, the argument I offer suffices, if it is succesful, to *rebut* the challenge offered to intuitionism - intuitionism is a stable position - without supplying a suasive argument which would satisfy a strict finitist who is prepared to adopt a scepticism about inferential practice. One might put the point by saying that *both* intuitionism and realism involve making some metaphysical assumption but that their respective assumptions are *distinct*, accepting the intuitionist's assumption has only been justified by an appeal to a general adequacy constraint on an account of a practice (viz., that the account allow for the practice of inference). But, because the realist assumption is at least as strong as the intuitionist's (since realism implies intuitionism, in terms of the gloss given on these claims at the outset of the chapter, i.e., that all sentences have determinate truth values entails that the truth values of all effectively decidable sentences are determinate[1]) I claim that there is an imperative to accept the intuitionist's assumption which the realist's metaphysical assumption lacks.

Let me sketch the form of my argument. My first claim is that anti-realist arguments lead to accepting a semantic theory which differs in *form* from a truth conditional account: truth conditions are no longer taken as fundamental but are explained in terms of verificatory or justificatory procedures. So naive realism must be rejected. This appears to leave three possible positions: ideal verificationism (explained below); intuitionism; and, strict finitism. I then argue that there is no slide from intuitionism to ideal verificationism, i.e., I argue that *if* we can be brought to accept an intuitionistic position we do not thereby commit ourselves to an ideal verificationist position. Finally I claim that the strict finitist position must be rejected since it fails to allow for an account of inferential practice. So this general constraint (of requiring an account of inferential practice) provides a motive for accepting an assumption of determinacy of truth value which goes beyond that which the strict finitist is prepared to allow. Since the intuitionist's assumption is distinct and is weaker than that of the ideal verificationist we can only claim that the general constraint provides support for the intuitionistic position.

It is worth distinguishing two sorts of realist; the first holds that we simply have a grasp of determinately obtaining, verification transcendent truth conditions for non-effectively decidable sentences. The second

---

1. This is not to say that the practice of intuitionistic mathematics must be *included in* the practice of realist mathematics.

tries to account for that grasp in terms of a grasp of infinitary verification procedures. Call the second form of realist an ideal verificationist. The parallel between the two arguments applies when we consider the ideal verificationist position since then the realist, although admitting that we have no effective decision procedure, insists that we have an infinitary decision procedure which justifies his belief in determinate truth values. That is his metaphysical assumption. Whereas the intuitionist assumes that possession of an effective decision procedure justifies belief in determinate truth values. That is *his* metaphysical assumption. So far the parallel seems true.

There is another aspect of the parallel. The intuitionist, in so far as he does not question our grasp of (presently) non-effectively decidable sentences, will not question the ideal verificationist's description of the infinitary procedure. What he questions is whether a procedure so described is genuinely a *decision* pocedure, that is, he questions whether possession of a description of such a procedure justifies belief in determinate truth values. Similarly the strict finitist (or, at least, the strict finitist whose position supposedly results from an anti-realist attack on intuitionism) will not question our grasp of an effective decision procedure for sentences involving, say, exponentiation. What he questions is whether that grasp justifies belief in determinate truth values for such sentences.

The parallel disintegrates once we turn to look at the accounts offered of our understanding of the relevant sentences. A classical truth conditional meaning theory is attacked by the anti-realist for its inability to give an informative account of what our understanding of verification transcendent truth conditions consists in (and thus cannot justify ascriptions of knowledge of such conditions). The anti-realist offers an alternative model of our understanding in terms of what counts as a justification for holding the sentence to be true. So our understanding of, say, universally quantified sentences consists in an ability to recognize of any construction whether it is an operation taking an arbitrary member of the domain onto a proof that the appropriate predicate applies to it. I take it that this aspect of the anti-realist's argument is endorsed by the strict finitist, intuitionist and ideal verificationist. Each position is distinctively characterized by what constructions it is prepared to recognize as encapsulating a proof or decision procedure. What this means is that each position accepts a proof theoretic explanation of the meanings of the logical constants. However because of the impredicativity of those explanations they establish, at most, a framework within which the meanings of the constants must be given. So there is still an issue as to whether the notion of proof used in those explanations is to be that of a classically valid proof (as the ideal verificationist contends) or intuitionistic.

A further point should be noted and that is that a proof theoretic explanation can only be anti-realistically acceptable if it remains a decidable matter as to whether a construction is to be accorded the status of a proof (of a given sentence). (Shortly, I shall consider an argument which implies that this view is fallacious.) Otherwise grasp of a given sentence cannot be taken to be manifested in an exercise of a recognitional skill. This rules out any *direct* appeal the ideal verificationist may make to infinitary verification procedures *as proofs*. At most the ideal verificationist can appeal to infinitary verification procedures in order to justify certain proof procedures.

Such an appeal is, however, illicit because it transcends anything that can be justified by appeal to the model of meaning which, according to the above, the ideal verificationist endorses, that is, a model of meaning which takes provability as central in the semantic theory. This, I claim, is a consequence of the incompleteness results. What these show is that grasp of the proof procedures relating to any sentence of the language cannot be taken to issue in a guaranteed means of decision even if the proof procedure is classical (justified by appeal to infinite verification procedures). The point here is that taking proof as central in the semantic theory means that we accept the orthodox intuitionistic explanation of the logical constants in terms of proof. There is a question as to what notion of proof is to be used in these explanations. The ideal verificationist takes himself to be entitled to assume a classical notion of proof. But then the incompleteness results show that even *classically* we cannot guarantee that there is a proof or disproof of every sentence. Thus one cannot appeal to classical provability in order to justify believing in determinate truth values for all sentences. So the appeal to infinite verification procedures in order to justify classical logic fails to justify belief in determinate truth values once the semantic theory is accepted to be of proof theoretic form. The end result is that there is an awkward mismatch between the justification provided by the semantic theory and the practice of inference.

§2.1 Sentence Size Bounds:

There is, though, an apparently strong ideal verificationist challenge to this position. The first stage of that challenge notes that the incompleteness results are accepted by intuitionists as showing that our notion of provability cannot be circumscribed within a single formal system. The ideal verificationist is able to make precisely the same move, that is, he can accept the incompleteness results as showing that his notion of proof transcends the limits of a given formal system.

The next and most crucial stage of the challenge questions the suggestion that proofs must be of finite length. That is, the ideal verificationist questions whether the status of a construction as a proof must be a decidable matter in the manner suggested by the intuitionist. This part of the argument can be drawn fairly directly from Langendoen and Postal's *The Vastness of Natural Languages* (1984). If both aspects of the challenge can be made good then the ideal verificationist can hold the we have a notion of (transfinite) proofs which is, i) not capturable within the resources of a single formal system, ii) sufficient to guarantee a determinate truth value for every sentence.

Langendoen and Postal attack what they see as an unfounded assumption common to a whole variety of linguistic theories, namely, that sentence size has no finite bound but that all sentences are of finite length (so sentence size is infinitely bounded). They consider arguments in favour of such views which attempt to show that any finite restriction on sentence size would be based upon "arbitrary" or not essentially linguistic factors (such as contingent performance limitations of human speakers). Such factors, it is claimed are not essential to the *grammaticality* of sentences. What Langendoen and Postal attempt to show is that such arguments are either question-begging or tell in favour of a rejection of *all* (including transfinite) bounds on sentence size; that grammaticality is in no way linked to size but is a purely formal property. Thus, they conclude, no restriction on the size of a grammatical sentence should be accepted.

Transposed to the setting of the controversy between the intuitionist and the ideal verificationist the argument would be that the intuitionist adheres to an ill-motivated finite size bound on the length of proofs/decision procedures. Any argument, it would be claimed, to show that we grasp arbitrarily long (but finite) decision procedures is either fallacious or shows that *any* bound on the length of proofs is arbitrary. In fact, Langendoen and Postal make just this point when considering an objection to their view which claims that in a constructive grammar a size law need not be explicitly formulated since a size bound will emerge as a consequence of the finiteness of derivations of sentences.

Let us look now at some of the detail of Langendoen and Postal's argument. First Langendoen and Postal attempt to point out the faults in certain common arguments against a finite size bound on sentences. Any argument, they claim, that is based on evidence of actual linguistic behaviour will not tell in favour of an infinite size bound since it will fail to discredit a finite but sufficiently large size bound. Also arguments based on the preservation of sentencehood upon repeated application of a given sub-sentential unit (e.g., "I know that ...") will fail for either of two reasons: i) such arguments beg the question over whether sentencehood *is* preserved when the application is repeated any (finite) number of times, i.e., such arguments assume that there is no finite size bound transgression of which invalidates a string as a sentence; ii) if the argument is construed as a *demonstration* that such strings are part of a natural language, then the demonstration will include an inductive step and so cannot be constructed within an axiomatic theory which does not include a question-begging axiom of infinity.

However Langendoen and Postal do think that a cogent argument can be made against any finite size bound on sentences. Any such bound they argue can be seen to lack a secure *linguistic* motivation. That is, any length property used to determine sentencehood will not be based on "the structural property of syntactic well-formedness" (1984, p.35 Note: Here Langendoen and Postal are quoting Katz, with approval) but will either be wholly arbitrary or motivated by contingent performance capacities exhibited by speakers. Langendoen and Postal attempt to show this by considering how one would construct rules governing grammaticality from an inductive base of actual attested sentences (which thus are necessarily finite). The argument is then as follows. Clearly we cannot take the maximum size of sentences in the inductive basis as a size bound for the language as a whole because we could easily imagine having an inductive base which included sentences of (far) greater length. In order to make this possibility vivid Langendoen and Postal describe the situation thus,

To see further that the principle [that sentence size is bounded by the maximum size of sentences in the inductive base] is a property of something distinct from L [the given natural language], imagine creatures, call them Woocoos, with life spans one million times greater than ours, with one billion times more memory and with comparable extra reasoning power. Clearly, the IB(L)s [inductive bases] of Woocoo linguists studying [natural languages] spoken by Woocoos would be subject to an entirely different and much weaker size principle than [the above]. Such principles thus provide no information about L but only about the process of linguistic research on L carried out by creatures with fixed limitations. (1984, p.37)

Alternatively if we choose some other larger finite sentence size bound then precisely the same imaginative example will again show that such a choice is arbitrary or is based on perceived limitations on performance. The point, they claim might be put as follows,

> There is no case where an intuition of ill-formedness is attributable to mere length. All that is ever observed is that as sentences become longer, they become harder to understand, perform, etc. (1984, p.37)

They go on to develop what they claim is a second argument based on similar considerations for the same conclusion. The point of the second argument is that the lawful properties of the language, i.e., the properties which distinguish the particular language, will be unrelated to sentence size. Thus importation of a sentence size bound will have the consequence of excluding infinitely many strings which satisfy the distinguishing properties for being a sentence of the language concerned and which thus *are* sentences of the language.

This second argument is important because Langendoen and Postal exploit the generality of the argument in order to show that *no* sentence size bound is legitimate. They note that since the lawful properties of the language are insensitive to sentence size *any* size bound will miscategorize certain perfectly well-formed strings as non-sentences.

This manner of proceeding is however muddled. The muddle emerges because Langendoen and Postal impute an unwarranted generality into the scope of the conclusion of the second argument. This in part arises because Langendoen and Postal are apt to treat the second argument as distinct from the first. It is not. Rather it is based wholly on the first argument. Langendoen and Postal note that their conclusion "follows from the assumption that P [the lawful properties of the language] are size-independent. But," they go on to note, "since the previous argument shows that there is no non-arbitrary linguistic basis for a finite length bound, this already follows." (1984, p.38) So without the support of the first argument the second argument only holds given an assumption which is both large and, in view of Langendoen and Postal's own earlier criticisms of other arguments, question-begging.

The argument against *any* size bound on sentences thus only goes through if an appropriate generalization of the first argument can be given. But this Langendoen and Postal do not attempt to give. Moreover it is hard to see how this step of the argument could be constructed. Recall that the first argument made use of the device of imagining creatures with finitely extended performance limitations. This device is not merely rhetorical. What in the transfinite case could possibly supply this step of the argument? Can we imagine beings with *no* finite limitations on performance?

The response then to Langendoen and Postal's argument is to say that in the transfinite case there is an appropriate size law (restricting sentences to finite length) and that this size law is not arbitrary and is not, at least in a restricted, purely formal sense of the term, linguistic. It *is* performance related but is not related to *contingent* limitations which humans actually possess. Rather it is related to what we can coherently imagine as exercise of skills constitutive of grasp of the appropriate linguistic rules.

Another way of putting the point is to say that *if* we can imagine beings which have appropriately infinite performance capacities then the first stage of Langendoen and Postal's argument in the transfinite case goes through. (Note that the argument must be repeated at each stage for each possible transfinite size law: no *general* argument against all size laws is immediately in the offing.) But to make this assumption is just to assume that our linguistic competence shows that we manifest a grasp of truth conditions, given by infinitary decision procedures, which determinately obtain irrespective of our ability to determine the relevant truth value. That however is *precisely* the conclusion for which the argument was intended to offer support. So Langendoen and Postal offer no non-circular vindication of the ideal verificationist's position.

§2.2 Infinite Decision Procedures:

Whether or not one agrees with Langendoen and Postal's argument for the lack of size bounds for *sentences* one might, plausibly, see fit to reject the extension of the argument to *decision procedures*. In this section I want to investigate reasons, drawn from considerations to do with the nature of decision procedures, for denying the coherence of infinite decision procedures.

I claim that a distinctive feature of a decision procedure is that one should be able to give a description of it such that at any point it is clear what the next correct step will be. Further, the description should include definite conditions for having completed the decision procedure. An infinite decision procedure fails to satisfy the second condition since either the completion condition is lacking or it is based on an absurdity (and thus is equivalent to a description lacking a completion condition). For example consider the following algorithm for a decision procedure for the sentence $(\forall x)Fx$ where F is decidable for each of its arguments.

1. Set $n=0$
2. Check Fn
3. If not-Fn then stop
4. If Fn then set $n=n+1$
5. Go to (2)

This algorithm lacks a full completion condition since the machine is instructed to stop only if not-Fn holds for some n. This patently cannot be guaranteed in advance. This, it might be thought could be remedied by using transfinite ordinals, i.e.,

1. Set $n=0$
2. Check Fn
3. If not-Fn then stop
4. If Fn then set $n=n+1$
5. If $n=\omega$ then stop
6. Go to (2)

This algorithm putatively has a stopping instruction in any eventuality. But since we know that $\omega \neq n+1$ for all values of n this halting condition is vacuous. This machine is thus (unsurprisingly) equivalent to the former.

Perhaps due to motives drawn from these sorts of consideration one often encounters a decision described as follows: Check $F0$ in the first minute; check $F1$ in the next half minute; check $F2$ in the next quarter of a minute; etc.. Now, provided we can conceive the possibility of implementing a procedure with an arbitrary but finite degree of rapidity, it would seem that we have described a decision procedure. Moreover it is a decision procedure which is guaranteed on the elapsing of two minutes to issue in a decision, so has a guaranteed completion condition.

We should however be alerted to the fact the something is awry by the need to tie completion of the procedure to this extraneous condition. Indeed if time failed to form an ordinally infinite uniformly dense set but was quantised then we would have failed to describe a decision procedure for the sentence $(\forall x)Fx$. Whether or not time *is* quantised may (for all I know) be a contingent matter or may be decidable *a priori*. The mathematical state of affairs would, in the former case, be a contingent matter. Whilst, if the latter were true, the resolution of a mathematical issue would depend upon (what I take to be) a deep result in the philosophy of time. This conclusion seems radically implausible.

So provided one grants my initial claim about the describability of a decision procedure no sense can be made of an infinite decision procedure.

Let me recap the course of my argument in this section. I claim to have revealed a disanology between the strict finitist attack on intuitionism and the intuitionist attack on realism. Two stages were involved in this demonstration. In the first I tried to show that accepting the anti-realist arguments must lead to adoption of a semantic theory which has a different form to that of the realist. That is, one is forced to adopt a notion other than truth as central in the semantic theory. Anti-realist semantic theories (accepted by both strict finitists and intuitionists) take a notion of provability (or decidability) as central. That position, it was noted, suffices to halt the slide from an intuitionistic rebuttal of strict finitism to classical truth conditional realism. But this simply opens up the possibility of a position, here called ideal verificationism, which accepts the form of an anti-realist semantic theory but questions the notion of proof or decidability which that theory takes itself to be entitled to. The thought here was that the ideal verificationist might argue that in just the same way as the intuitionist rejects tying our grasp of decision procedures to *actual* performance limitations, one should reject a similar constraint based on implementability *in principle*, or to *finite* performance limitations. I then went on to argue against the notion of infinite decision procedures. The upshot of this discussion was that the (metaphysical) assumption motivating ideal verificationism is *distinct* from that motivating intuitionism. So no slide from the latter position to the former has been demonstrated. What I want to do now is to look more closely at the debate between the strict finitist and intuitionist in order both to clarify and to assemble some motivation for the latter position. To do this we must look to the character of our understanding of the relevant logical and non-logical vocabulary. Since we are concerned with when we are justified in holding sentences of the form $F(\underline{n}) \vee \sim F(\underline{n})$ to be true, I shall be interested in our account of negation and disjunction and in our grasp of predication.

§3 Implications of the Compositionality of Meaning:

In this section I am assuming as a result from the previous section that the anti-realist's notion of decidability is not to be explained in terms of infinitary decision procedures. Given that, it is then true to say that, since there is no way of uniformly extending our grasp of quantification over arbitrary finite domains to that over infinite domains, quantifying over infinite domains is problematic because here we have a *specific* operation which introduces undecidability. Understanding of the quantification-free region of language consists in a sensitivity to the possession and correctness of proofs. The incompleteness results show that once we introduce quantification into the system of quantification-free arithmetic to create first order arithmetic then there will be sentences which are perfectly understandable but for which those proof methods, sensitivity to which is constitutive of the ability to understand the quantification-free language, do not suffice to establish a means of decision. Thus we cannot presume that these sentences have determinate truth values, i.e., there can be nothing in the understanding of the quantification-free language which shows that in grasping quantified sentences we grasp truth conditions which, although transcending determination, determinately either do or do not obtain.

In contrast, in the case of unsurveyably large finite procedures no *specific* operation is introduced which accounts for the indeterminability in practice of truth values. Typically, we construct sentences which are not, in practice, decidable by combining expressions denoting operations which, in the right context, have surveyable applications. Thus it is only when we apply the predicate "is prime" to large enough, though surveyably expressible, numbers that we develop practically undecidable sentences. Now it would seem that this position should be problematic only to a strict finitist who insisted that we did not understand such sentences and the strict finitist need not dispute this. What he disputes is simply the assumption that such sentences have determinate truth values.

However this ignores the rôle of the decision procedure. The realist might attempt to buttress his position by appeal to considerations to do with the compositionality of meaning. That is, he might try to claim that our understanding of the components of an undecidable sentence ensures that we grasp a truth condition which determinately obtains. The intuitionist objects to this precisely because he takes quantification over infinite domains to be an operation which gives rise to certain sentences which resist determination by presently understood decision procedures and, until we can guarantee a means of dissolving this resistance, we cannot assume determinacy of truth values. But, in the cases that the strict finitist objects to, as a matter of our contingent performance limitations, it comes about that an attempt to institute the *same* procedure which we successfully implement in small finite cases becomes impossible in large finite cases. Grasp of the procedure in the former set of cases manifests grasp of a procedure which issues in determinate truth values. Since this is the same procedure in the latter set of cases it must, even when implementation is frustrated (by our own feebleness), issue in determinate truth values.

Notice the way this argument is intended to work. I am not simply saying that in all cases we have grasp of a procedure such that the following, on occasion counterfactual, claim always holds good; if the procedure is implemented it determines a truth value. Rather I am saying that where it is feasible to implement the procedure the intuitionist and the strict finitist take the mere *possession* (rather than implementation) of such a procedure as a justification of our belief in determinate truth values. Now since it is the *same* procedure invoked in the unsurveyable set of cases we are again able, according to the intuitionist, to use it to justify our belief in determinate truth values irrespective of the actual implementation of the procedure. The point is that assuming a relevant sameness in our grasp of the decision procedure across the surveyable/unsurveyable boundary justifies the move from taking it, uncontroversially, that (i) implementing the procedure issues in a determination of truth values to taking it that, (ii) it is determinate which truth values the procedure will, in any case, issue in. We need now to look more closely at what this "sameness" in grasp of the application of a procedure consists in.

§4 Understanding the Application of a Predicate:

Assume that we have a predicate, F, for which we have a well defined criterion of application, satisfaction of which for any argument is an (intuitionistically) decidable matter. Possession of such a procedure associated with a given predicate is taken by the strict finitist to justify assertion of $F(\underline{n}) \vee \neg F(\underline{n})$ only when we can actually implement the procedure for the given argument. The intuitionistic challenge based on the compositionality of meaning asks what in our understanding of any of the components of this disjunction distinguishes those instances when it is (strict finitistically) assertible from those when it is not.

Let us first focus on the understanding of the predicate. Grasp of the predicate is (at least partly) constituted by mastery of the criterion of application for the predicate. There are two ways in which a strict finitist might reject the assertibility of $F(\underline{n}) \vee \neg F(\underline{n})$ when we cannot, in practice, determine satisfaction or failure to satisfy the criterion. One sort of strict finitist might reject the *intelligibility* of the disjunction, claiming that for sufficiently large $\underline{n}$ we have no grasp of $F(\underline{n})$. This strict finitist questions our grasp of certain contents .

The second sort of strict finitist would accept that we grasp the disjunction but would go on to claim that, since we cannot in practice implement the relevant procedure, we are not justified in believing that the sentence has a determinate truth value and thus are not justified in asserting that LEM holds for it.

I have repeatedly remarked that the first sort of strict finitist is not a position which results from accepting distinctively *anti-realist* arguments. Those arguments are not revisionary about content, that is, those arguments do not issue in a challenge to our grasp of given sentences. We certainly understand the expression "$F(\underline{n})$" since we understand the component expressions and their mode of composition. The anti-realist then asks what that understanding consists in. His point is that it cannot be taken to consist in grasp of bivalent truth conditions for sentences whose truth values we cannot guarantee a means of decision. We cannot manifest such a grasp. Rather, what we manifest, in grapsing the components of the expression (and their mode of composition), is a sensitivity to the status of constructions as proofs or refutuations of the sentence concerned. This grasp will not justify believing that the truth value of the sentence is determinate. So for the anti-realist there is no question of whether or not we grasp "$F(\underline{n})$", there is only the question of whether that grasp justifies believing in determinate truth values.

Thus, if strict finitism is to be a consequence of anti-realist arguments it must be a strict finitism of the second sort. But in this case the strict finitist and the intuitionist agree in the account of our understanding of the relevant (elementary) predicates. The strict finitist simply questions whether that understanding justifies belief in determinate truth values when the procedure is only implementable in principle. This, however, means that the strict finitist and the intuitionist differ in their account of the meaning of vel.

§5 The Meaning of Vel:

So let us concentrate on the strict finitist who agrees that we manifest grasp of the application of a procedure to any number, i.e., that we understand the sentence "$F(\underline{n})$", for any n. In this case the strict finitist would not object to the intuitionist use of the compositionality of meaning but will claim that the meaning of "v" will not warrant the assertion of $F(\underline{n}) \vee \sim F(\underline{n})$. I want, though, to argue that in this case the strict finitist allows just the appeal to meaning which, combined with the need to accept the possiblity of inference, enables us to conclude that the truth value of each instance is determinate; that our understanding of, say, the sieve of Eratosthenes justifies us in holding that each number either is or is not prime. Now presumably even if the strict finitist is to follow us this far in the description of his position he will simply point out that, although the description of what we grasp may be correct, the notion of understanding is insufficiently robust to support the appeal the intuitionist needs to make to it.

The only response I can recommend to the intuitionist here is to reflect the challenge back on the strict finitist. Given that the latter accepts that our understanding of the procedure is, in appropriate respects, the same in the transition from the surveyable to the

McDowell takes from Dummett's (1978b). There Dummett notes that to have the concept square is to be able to discriminate between square things and things of other sorts. One of the ways of doing this is to apply the word "square" to recognizably square things. Two points should be noted. First, one could, prelinguistically, have the concept square but having the concept is not a *precondition* for learning the meaning of "square". Rather learning the correct use of the word "square" is one way of grasping the concept square. The point here is that we do not use the idea of a prelinguistic grasp of the concept to explain how meaning is imbued in language, i.e., how symbols come to have a representational rôle, rather, the possibility of a prelinguistic grasp of the concept entitles us to claim that the ability we are ascribing does not depend upon any specifically linguistic achievement. One could say that language is meaningful only because behaviour is. The theory of meaning gives an account only of linguistic meaning. Secondly, what one manifests in grasping the concept *square* is not a regularity in behaving differently relative to square things but a *discriminatory* ability, that is, a sensitivity to success or failure in this practice. This sensitivty to correctness (i.e., this element of rationality) is not essentially linguistic so, although fundamental to meaningful use of language need not be part of a complete description of what it is to use a language. Thus we can give an account of language which grounds possession of linguistically expressible concepts in non-linguistic but rational practices. Moreover, in doing so, we do not explain meaningful use in terms of antecedently grasped concepts.

This concern with the distinction between a regularity in behaviour and behaviour or use which is subject to standards of correctness readily calls to mind Wittgenstein's preoccupation with what it is to follow a rule in *Philosophical Investigations* and *Remarks on the Foundations of Mathematics*. McDowell's reservations about full-bloodedness seem to stem from his reading of Wittgenstein. In the example to suit my ends. Naturally, I do not think that this is so. Let me briefly explain the underlying thought. If we are to imagine a genuinely foreign community then we must make sense of an interpretative exercise. We cannot simply assume that our understanding of numerical terms agrees with theirs without unpacking that assumption. If the assumption is based on a coincidence of counting practices then the example asks us to imagine a single community where counting practice leads to a contradictory inferential practice. If not, then we must interpret their counting practice. We would then want some explanation of that practice (this is especially true if the community held absurd arithmetical beliefs such as 29+17=45). That is, we would expect them to offer or to recognize something like the Peano axioms as explanations of their practice. Once they did this our grounds for rejecting their practice would be that it is demonstrably inconsistent. This may well raise further questions about methods of demonstration but such questions of the soundness of the underlying logic are just what concern us. My contention is that such questions can be legitimately raised by the anti-realist provided the soundness of the one system does not imply the unsoundness of the other. The point seems to be that the possibility of raising the dispute requires a measure of agreement, not that the dispute cannot be raised.

If this argument holds then we have no argument for the view that the anti-realist cannot concern himself with questions of soundness. So it is open to the anti-realist to question whether classical logic can be shown to be sound (even though he admits that it is not unsound). The point is that classical logic can be shown to be sound relative to truth as realistically conceived, and cannot be shown to be unsound relative to the mere existence of non-inferential warrants. (The latter follows from the former and from the specifically anti-realist considerations advanced by Wright.) The question though is whether we are justified in ascribing grasp of realist truth to ourselves. The anti-realist gives a unsurveyable it is up to him to delineate a relevant difference in these two contexts of application. The "relevant" difference must be a *semantic* difference, i.e., it should not merely be a consequence of our contingent performance limitations. To put the matter slightly differently, the intuitionist claims that we grasp the procedure in such a way as to have a quite general understanding of its correct application. In each instance, in which the procedure is relevant, we take this understanding to justify a belief in the determinacy of truth values[1]. The strict finitist accepts the first but rejects the second assertion. He can do so either by denying that the understanding described in the first assertion ever, on its own, justifies a belief in determinacy of truth values or by showing how the justification fails in passing from the surveyable to the unsurveyable. But according to the first alternative possession of an actually implementable procedure will not justify belief in determinate truth values, only actual implementation of the procedure will suffice. But then it is utterly mysterious as to what possible use the strict finitist might have for reasoning by dilemma. Thus it would seem all but the most extreme finitist must reject the first alternative. What possible reason can there be for accepting the second given that the transition marks no mathematically significant boundary, that is no boundary which can be crossed only by acquisition of an essentially new recognitional skill? This notion of "an essentially new recognitional skill" is unclear. We could motivate the notion by showing that any boundary between actually implementable and in principle implementable procedures would be altered were our performance limitations to be finitely extended.

The problem here is that, given agreement on what an understanding

---

1. This situation, it should be noted, is appropriately disanalogous to that in which the realist finds himself since the realist is in possession of a sentence, which the intuitionist will grant is of definite sense, but which has no mutually understood decision procedure to justify his belief in determinate truth values, however the intuitionist's decision procedure is understood by the strict finitist. (See §2 of this chapter.)

of the predicate consists in, the intuitionist is simply baffled as to whatthe strict finitist takes the meaning of "v" to be. Now, if the strict finitist claims that being in a position to assert $A\mathrm{v}B$ is to possess a proof of either A or of B then the intuitionist can understand his use of "v" but would wonder why the possession of an actually implementable decision procedure ever warrants assertion of a disjunction. If the finitist is hard nosed enough to deny that then it is difficult, from either perspective, to see what possible use a finitist might have for "v" (unless he knew he had proved one of the disjuncts but had forgotten which). (This corresponds to the first alternative of the previous paragraph, i.e., rejection of the idea that mere *possession* as opposed to *implementation* of an actually implementable procedure justifies belief in determinate truth values.)

Let us assume then that the strict finitist notion of "v" is given by the following stipulation: $A\mathrm{v}B$ is assertible when we have a procedure which we can *actually* implement to give a proof either of A or of B. One way of putting the intuitionist's point would then be to make use of the strict finitist's admission that the extent of our actual recognitional capacities constitutes a contingent fact about us. The intuitionist could then show that either the finitist's stipulation fails to confer a determinate sense on "v" because the notion of actual implementability depends on empirical determination or that any precisification, in mathematical terms, of that notion enables the intuitionist to give a finitistically acceptable account of "v" which broadens the conditions under which a disjunction is assertible. The finitist's understanding of "v" becomes unstable since perpetually subject to extension. (The extensibility of the notion is not, in my view, problematic in itself since I think that some of our concepts are subject to continual extension. But I do think that the possibility of extension militates against finitistic restrictions of the notion and that if the idea of *extending* as opposed to changing the concept is to make sense then we have to have a stipulation whose form encompasses the process of extension. An intuitionistic stipulation allows this possibility. See chapter 6 for more discussion of this.) So, given an account of actual implementability, there would be borderline cases of surveyability for which the intuitionist can explain to the strict finitist how a small finite extension (small enough to guarantee that both can have access to notions of this size) of our actual capacities would enable us surveyably to implement the procedure. So if "v" was explained in terms of this small extension of our actual capacities the strict finitist would have to accept the assertibility of the disjunction. Thus in any attempt to draw a boundary by mathematical criteria alone the intuitionist can force the strict finitist, by considering only the contingency of our recognitional capacities, either to expand the boundary or to accept that the boundary is not *semantically* relevant..

But the dispute is difficult to apprehend in these terms. Let us again consider the transition from $A \to (B \vee C)$ to $(A \to B) \vee (A \to C)$, i.e., the passage from my (i) to my (ii) above (the end of §3), which Wright (Wright 1980, pp.208–9) points out is, in general, not intuitionistically acceptable. Wright, in *Strict Finitism* (Wright 1982) gives us the detail of an argument which purports to justify the transition in the present (mathematical) context without contravening any intuitionistic scruples. But he goes on to criticize the argument for relying on assumptions about the objectivity of meaning. Wright notes that although, in general, the passage from sentences of the form of (i), i.e., $A \to (B \vee C)$, to sentences of the form (ii), i.e., $(A \to B) \vee (A \to C)$, is intuitionistically invalid, he lists four conditions satisfaction of which justifies the inference[1]. So with these four conditions in place we are justified in accepting the transition. But Wright notes that the transition assumes that we can so confer meaning on our language that it is determinate what result correct implementation of a procedure should issue in and that this must hold both in advance of implementing it and irrespective of whether or not we actually can implement it. Put simply, Wright argues that where A hypothesizes the implementation of a decision procedure and D is provable from the disjunction of B and C we can, simply enough, prove $A \rightarrow D$ and the possibility of implementing the procedure hypothesized by A then allows us unconditionally to assert D. (Since $(B \vee C) \rightarrow [(A \rightarrow B) \vee (A \rightarrow C)]$ we can then assert the disjunction of conditionals, i.e., we justify the transition just mentioned.) In his argument Wright concentrates on the case where C is $\neg B$, that is, where A hypothesises the implementation of a decision procedure for the sentence B. The problem Wright discerns in the case he gives for the unconditional assertion of D is that in accepting it we implicitly accept the investigation independence of D. Only if we regard D as standing for a state of affairs which exists independently of our actually coming to know it can we take it that actual implementation of the procedure, A, is irrelevant to whether or not D, in fact, obtains. But, Wright continues, the conditional $A \rightarrow D$ simply means that we can extend our state of information from one of having implemented A to recognizing the truth of D. Knowing that we can

---

1. The first condition is that the hypothesis holds; that satisfaction of A brings about the truth of the disjunction. The second is that there is no indeterminacy as to which disjunct is true when the antecedent is satisfied. The third is that there is no extra condition such that its satisfaction in conjunction with that of the antecedent brings about the truth of one disjunct while its failure in the same circumstances brings about the truth of the other disjunct. Finally that the truth of the disjunction is in fact linked to the truth of the antecedent. So it must not be the case that, whether or not A holds, there are grounds for expecting the truth of $B \vee C$.

implement A does not then justify us in unconditionally asserting D since it may be that implementation of the procedure *brings about* the truth of D. The obverse of this position is that if we accept the argument as establishing the unconditional assertibility of D in these circumstances then we assume that, in advance of applying and irrespective of whether we ever actually apply the decision procedure, D has a determinate truth value. But this betrays a belief that we have succeeded in conferring a meaning on our terms such that fidelity to that meaning establishes standards of correct use which transcend our actual use. Note that there is a sense in which this objection is irrelevant to the issue of the intuitionist's use of procedures implementable only in principle since the objection would hold equally against a strict finitist who held that mere possession of an actually implementable procedure justified belief in determinate truth values. It is still not clear in that case whether the implementation of the procedure brings about the truth of D or discovers D to be true.

Wright's formulation of the central notions is unhelpful. That it is so can be seen from the way he resurrects the metaphysical mode of debate. The question that so fascinates Wright concerns what he terms the investigation independence of certain states of affairs. This finally comes to rest on questions about whether or not, in extending our state of information, we bring about (the truth of) those states of affairs. But this whole set of concerns relates purely to the notion of truth we take ourselves to be justified in using, that is, whether or not we regard truth values as determinate when they transcend recognition.

Obviously we can replace D by $B \vee C$. The intuitionist regards himself as justified in unconditionally asserting this disjunction so the play Wright makes with the transition between the conditional with disjoined consequent and the disjunction of conditionals is irrelevant provided we can simply justify the assertion of $B \vee C$ unconditionally. Moreover since it is a question of precisely this form which Wright takes to determine the validity of the transition the question of the validity of the transition must be, at most, secondary. Wright's formal manœuvres cloud the central issue which is the meaning of intuitionistic disjunction.

Wright indeed raises this issue in criticizing Heyting's informal explication of disjunction; that a proof of $B \vee C$ is a proof either of B or of C. If we are to take this as setting the meaning of "v" then the intuitionist can have no use for it in his deductive practice since any proof proceeding through the disjunction could, in practice, proceed through the proof of one of the disjuncts. So the meaning of "v" needs some liberalization. Wright takes it that the conditions warranting the assertion of a disjunction should be expanded so as to include those conditions in which the assertor can recognize that he is in a position in which he can find out, expand his state of information to include, the truth of one of the disjuncts even if, at present, he cannot say which. But this way of stating the proposed emendation of Heyting's explanation calls into question what expansions of one's state of information are possible and whether the truth so discovered had a status independent of the actual process of expanding one's state of information.

Instead consider this alternative explanation; a proof of $B \vee C$ is any construction (together with a method of applying it) of which it can be recognized that, if appropriately applied, it yields a proof of B or of C. Now the conditions under which the disjunction is unconditionally warrantedly assertible result just from the recognition that we have a decision procedure hypothesized by A such that $A \rightarrow (B \vee C)$ (which is simply the conditional all sides agreed we were capable of recognizing as true purely as a result of understanding the decision procedure). We stipulate the meaning so that it does not depend on recognizing conditions into which we can, in some sense, get ourselves, rather it depends solely on conditions which we can recognize ourselves to be in at present. We can now justify the assertion of the disjunction as a consequence of the meaning of "v". The one point at which a strict finitist might differ from the intuitionist could only now relate to when the former took himself to be in possession of such a general decision procedure. But, as just remarked, the possession of such a general procedure was agreed upon as a premise to the argument.

One question remains. Does not this new definition justify the, generally intuitionistically invalid, transition from $A \to (B \vee C)$ to $(A \to B) \vee (A \to C)$? Quite obviously it does no such thing since A must be something whose truth we can guarantee to bring about. But even that way of stating the position is confused. The meaning stipulation tells us when we can regard a certain construction as a proof. The construction given in the stipulation is an operation from constructions (we cannot be more specific than this, it may act on proofs or perhaps elements of a domain) to proofs so the sort of proposition which can take the place of A will be severely limited.

We should bear in mind that this argument, even if succesful, seems only to achieve a local justification of intuitionistic logic for mathematical discourse. Although I think there are important difficulties in extending this position to empirical discourse I do not think that these problems are to do with the internal coherence of an anti-realistically motivated intuitionism. Rather the problem is one of making good the notion of a decision procedure when transferred from the narrowly mathematical setting to the more general empirical one. We have firstly to give some account of conclusive but defeasible (i.e., criterial) warrants and then give an account of allowable or canonical ways of achieving these, i.e., ways which count simply as observing or measuring that the warrant is available and those which alter the conditions of assertion themselves. But, here, I cannot even speculate on the feasibility of this large programme.

§6 The Conservativeness of Intuitionistic Logic:

The stipulation, however, does not, of itself, solve the problem. We need to justify it. The realist, we know, claims to be able to characterize conditions (namely, any at all) under which the assertion of a disjunction of a sentence and its negation is assertible. But this commits him to holding that one or other is true irrespectively of whether we can determine the matter. Similarly, the intuitionist has characterized conditions under which he holds a disjunction to be true and this commits him (since he will endorse the classical truth-functional account as being correct although not explanatory) to holding that one of the disjuncts is determinately true or false in these conditions irrespectively of whether we actually can determine or have determined this.

What we need to look at is how, given their favoured meaning stipulations, the intuitionist justifies applying excluded middle to decidable sentences and the realist to any sentence at all. Take the latter first. The condition in which $A \vee \sim A$ is true is just the condition (via the meaning stipulation for "v") that A is true or that $\sim A$ is true and that condition in turn is just the condition (via the meaning stipulation for "$\sim$") that A is true or that A is not true (i.e., A is false). But, clearly, we cannot use this explanation of negation and disjunction to justify LEM since holding that the final condition so characterized is universally true is just to hold that LEM holds for all sentences of the form "A is true" and this, given the equivalence principle commits us to LEM itself. So as an attempt to justify determinacy of truth values the process is circular.

Now let us turn to the intuitionistic explanation of $A \vee \sim A$. For the intuitionist a proof of $A \vee \sim A$ is any construction of which it can be recognized that, if appropriately applied, it yields a proof either of A or of $\sim A$, i.e., it yields a proof either of A or a construction of which it can be recognized that applied to a proof of A it yields absurdity. The problem for the intuitionist is that this definition of the meaning of "v" does not give a conservative extension of the base class of statements relative to recognition of truth, i.e., relative to possession of a proof.

We need to explore a little further the consequences of the stipulation. Most importantly we need to discover what property of sentences is preserved in inferences. Consider a derivation of C which proceeds by a step of vel elimination on $A \vee B$, i.e., we have a proof of $A \vee B$, and proofs of C from the assumption of a proof of A and B respectively. According to the stipulation this means that we have a construction which can be recognized to issue in a proof of A or of B and constructions, recognizably, which take proofs of each of A and B into proofs of C. So we can form a construction which, if appropriately applied, recognizably issues in a proof of C by concatenating the above constructions in this way: apply the construction which issues in a proof of A or of B, if the outcome is a proof of A apply the construction which transforms this into a proof of C and similarly if the outcome is a proof of B. Note that we are not begging any questions by assuming that the description just given succeeds in defining a legitimate construction; we have not assumed that it is determinate which of A or B will be proved when and if the construction is applied. We assume that our understanding of the construction guarantees that no other outcome aside from a proof of either A or B is possible and that in either case we can transform the outcome into a proof of C. (Essentially we are only repeating here the idea that if we understand a procedure then even if we cannot implement it we know that implementation of it would issue in a determinate conclusion.) Now this simple sketch suggests that the logic applies conservatively relative to the notion of having a construction which recognizably issues in a proof of the sentence concerned.

§7 The Relation of Intuitionistic Proof to Intuitionistic Truth:

What is the relation of this notion of having a construction which issues in a proof to truth? We cannot equate these two notions without relinquishing the idea that truth distributes across the disjunction since a proof of AvB simply requires that we possess a construction which issues in a proof of one or the other and this is weaker than having a construction which issues in a proof of one or having a construction which issues in a proof of the other. But the two notions do coincide when we are only considering the assertion of the sentence as a whole. They only come apart when we consider sentences as constituents of other sentences. We learn from the behaviour of sentences in complex sentences that truth is a wider notion than possession of a construction issuing in a proof, i.e., we can infer from the behaviour of sentences in disjunctions that a sentence may be true although we do not possess a suitable construction. It becomes difficult to characterize precisely in terms of proof what truth is. But that need not concern us since we rely on the equivalence principle and the legitimacy of a truth functional description of the logical constants to account for truth.

So, having characterized the metaphysical dispute in terms of a belief in determinacy of truth values we can settle the question of the appropriate notion of truth by asking when the law of excluded middle is applicable. The, unsurprising, answer is that we can assert $Av{\sim}A$ just when we have a decision procedure for A which is effective in principle.

The question now is, does a semantic theory characterizing proof conditions give an adequate account of meaning? The thought here is that in characterizing what counts as a canonical proof of D, say, we have given a complete account of its assertibility conditions and thus of its meaning. But now we are asked to acknowledge that its assertibility conditions are wider since it may be asserted as the conclusion of a "proof" proceeding via a disjunction neither of whose disjuncts has been proved. So either we have not succeeded in characterizing its canonical proof conditions or its canonical proof conditions do not exhaust its assertibility conditions (it is a matter of choice how we describe the situation).

We have now three notions in play: canonical proof; constructions which recognizably issue in canonical proofs; and truth. The meaning of a sentence is determined by what counts as a canonical proof of it. Our logic applies conservatively relative to the notion of having a construction which recognizably issues in a (canonical) proof. A sentence may be true although we do not possess a construction which recognizably issues in a (canonical) proof of it. What I want to show is that the relation between canonical proof and constructions which recognizably issue in canonical proofs allows for a molecular theory of meaning and that the conservativeness of our logic relative to the latter notion establishes a sufficient epistemic constraint on the notion of truth.

We admit that proof conditions transcend canonical proof conditions and that truth transcends proof. The situation thus bears a profound resemblance to that described by Dummett in *The Justification of Deduction* where he notes that in order to account for the usefulness of deduction we have to acknowledge some gap between truth and its recognition by direct or canonical means (Dummett 1975a, p.314). But if canonical proof conditions are meaning determining then surely the semantic theory should apply conservatively relative to the existence of canonical proofs (in which case we have to justify holding that the non-canonical proof gives us the right to assert the existence of a canonical proof because, say, it shows how the latter could, in principle, be constructed - this does close the circle of explanation). Whilst, if they do not then we have introduced a degree of holism into our theory; the meaning of sentences does, now, depend on that of more complex sentences since constructions issuing in canonical proofs may proceed via more complex sentences. Is that acceptable?

The intuitionist must avoid impaling himself on either horn of this dilemma. He must, it seems to me, hold that his theory is not holistic and thus that canonical proof conditions are meaning determining but that the notion of conservative extension need not apply to the existence of canonical proofs but to his notion of proof as appropriately explained. This leaves him in need of an account of the relation between his notion of proof and canonical proof. Here I have no penetrating insight to offer but think that none such is needed. The relation, I claim, is transparently given in the account of proof in terms of recognizing of a construction that if it is appropriately applied it yields a canonical proof. This account of proof gives a general satisfaction condition for constructions to be proofs once the canonical proof conditions are known. Having this account enables us to claim that canonical proof is still determinative of meaning since the assertibility conditions of the sentence flow from its canonical assertibility conditions. So I rest my case here and await a challenge to the coherence of holding that meaning is determined by canonical proofs, that proof is as just explained and that the semantic theory applies conservatively relative to this notion of proof.

The realist, too, claims that his semantic theory for logic shows logic to apply conservatively relative to his notion of truth. So let us be sure about the differences between the realist and our intuitionist. The realist, firstly, takes truth to be a semantic primitive so it is difficult to see how he *justifies* the bivalent behaviour of his notion of truth. In contrast the intuitionist gives a contentful account of his notion of proof and explains the manner in which it applies to sentences in terms of this account. The canonical proof relation (i.e., the relation of a construction to a sentence) is always decidable. Thus the semantic theory gives a suitably reductive account of our grasp of undecidable sentences since it never calls for us to admit that the truth value of a sentence might determinately transcend our capacities to determine it in principle. Our grasp of any sentence consists in being sensitive to the status of constructions as canonical proofs or refutations of it and this status is always a decidable matter. Evidently the proof relation is also always decidable so we encounter no bar to our explanation of undecidable sentences. (Although obviously, in both cases, the existence of a proof is not always a decidable matter.) The applicability of the truth predicate, in contrast, is not always decidable and thus not only is the direct means of recognition of truth, on a realistic interpretation, often inaccessible to us but, since meaning for the realist consists in grasp of bivalent truth conditions, it is utterly mysterious how any sensitivity we show to verificatory procedures relates to the understanding of sentences whose truth values are undecidable. Our intuitionist avoids entirely this problem which is endemic to realism. The reason for this is that the intuitionist's notion of truth may transcend verification but its obtaining is always linked to the satisfaction of general proof procedures and these, in turn, are explained relative to canonical proof procedures which are meaning determining. So although his notion of truth transcends recognition by direct means it is not transcendent with respect to all verificatory procedures. These differences pave the way for the establishment of a coherent intuitionistic position (at least in repect of mathematical discourse).

To sum up, the intuitionist gives an account of meaning by recusively characterizing canonical proof conditions. His semantic theory gives an account of logic which applies conservatively with respect to his notion of general proof conditions which, in turn, receive explication in terms of the notion of canonical proof. The notion of truth is explicated via the equivalence principle. The intuitionist is thus able both to give a suitably molecular semantic theory which obeys his constraints on an account of meaning and, in so doing, he acquits himself of his obligation to give a communicable account of his notion of truth.

The point of the above argument was to show that grasp of the general applicability of a decision procedure justifies the belief in determinacy of truth values when we can make sense of the idea of applying that procedure. Although considering cases where that procedure cannot in practice be implemented I have not committed the strict finitist to the intuitionist's notion of truth. If the argument is successful it only succeeds in questioning the strict finitist's motivation. The reason for this is that the question now reduces to one of what the strict finitist might mean by "any number", in particular, does his restriction of that notion guarantee that the procedure is, for each number, implementable in practice? Surely not, since provided our method of representing the numbers is not too complex we shall be able to give instances where, although we can surveyably represent the number, we cannot surveyably implement the decision procedure. Thus our strict finitist is forced to admit the determinacy of truth values for those instances where we have only a decision procedure which is effective in principle. His position thus collapses: we have no reason to suppose that his conception of the natural numbers differs from that of the intuitionist.

§8 Strict Finitism Relies on an Interpretation of Rule Following which is not Obligatory to the Anti-Realist:

Another important difference between the intuitionist and the classicist is that grasp of the intuitionist's notion of proof, unlike the classicist's notion of truth, can be manifested in our understanding of logically simple sentences. This means that it is appropriate to justify our logic in terms of it being a conservative extension of the base class relative to this notion of proof. For example the fundamental theorem of arithmetic states that every number has a unique prime factorization. We prove it by showing that it is true of an arbitrary number. Now we can only recognize this as a proof if we can recognize that we have a construction which applied to any number yields its prime factorization. Admittedly the predicate here is not simple but complex. But the idea is clear: we can, on occasion, recognize that a procedure could be implemented although we may not, in practice, be able to implement it. I have tried to argue that this position is coherent from an anti-realist point of view. It is conceivable that Wright would object even to this on the grounds that in believing that the procedure has a determinate application in cases when we cannot apply it we betray a commitment to the objectivity of meaning. I think that this accurately reflects Wright's position. What I want to do now is to show that an objection along these lines depends upon Wright's view of the applicability of the Rule Following Considerations. In the last chapter I tried to argue that this view is not mandatory for the anti-realist in which case it is not clear that there is an argument from general *anti-realist* grounds to strict finitism (as opposed to intuitionism). Lastly I want to suggest that the interpretation of the Rule Following Considerations that I sketched in the last chapter lends support to intuitionism.

The intuitionist just described relies on the thought that if we understand the domain of numbers and have an understanding of a procedure applicable to numbers then we have an understanding of the application of that procedure to any number. Moreover he supposes that that understanding justifies us in asserting what the results of implementing that procedure would be were we to implement it.

Wright's objection would be that the position presupposes that we have conferred a meaning on our terms such that using them correctly in accordance with that meaning requires a certain pattern of use which outruns that which we are able to make. So the picture we are given of what understanding consists in is of grasp of a determinate pattern of use which cannot be displayed in the finite set of instances of use which we make of the term nor is the pattern of use evident to the learner in his exposure to the use others make of the term. Such a notion of understanding is comprehensively demolished in the Rule Following Considerations and is one the anti-realist can have no truck with. So much then for an outline of what I take would be the brunt of Wright's objection to the above position.

The outline suffices to motivate the thought that Wright's reservations about intuitionism are motivated by his view of the Rule Following Considerations as providing a critique of the objectivity of meaning. So, at the moment, we have an argument based on Wright's interpretation of the Rule Following Considerations against the intuitionist position. I want now to go on to show that an alternative interpretation of the Rule Following Considerations does not issue in an attack on intuitionism. This, provided the interpretation I give is consistent with anti-realism (this is argued in the previous chapter), will show that Wright does not succeed in outlining an *anti-realist* reason for dismantling intuitionism. It would be nice also to be able to show that all strict finitist attacks must depend upon reservations about meaning which parallel those of Wright since then I would have a comprehensive defence of an intuitionism resulting from anti-realist principles. However I know of no argument to demonstrate this. But it does seem clear that the only appeal the intuitionist wants to make is to the nature of our understanding and that an objection to that would have to parallel Wright (even if the motivation for the objection was different in that, say, it was grounded in a specific ontological view) at least as far as showing that the character of our understanding was insufficiently robust to support such an appeal. Now if I can show that an anti-realist position will allow such an appeal then I will have shown that, pending an argument to show that this anti-realism misdescribes the character of our understanding, intuitionism is a valid position.

§9 Intuitionism Defended:

So now I must defend my brand of intutionism. The broad anti-realist position I have been advocating in the last two chapters seeks to ground linguistic practice in more general human practices which themselves are subject to normative constraints but which are not susceptible to philosophical explanation. This position allows us to describe understanding of certain basic expressions not as grasping a rule or interpretation but as a way of acting. So Rule Following Considerations gain no grip on the understanding of these expressions. From this perspective there is no temptation to limit our grasp of correct use to use against a securable background of communal assent. The rôle of the community in this account is in providing a normatively constrained practice in the context of which our actions and behaviour have meaning for other practitioners. The important distinction here is that between a community whose members display a shared regularity in behaviour and a community whose members understand one another (even though that understanding depends upon there being sufficient resemblance and regularity in their behaviour). The position has been clearly mapped out by McDowell (McDowell 1984)[1]. McDowell claims that it is an essential aspect of meaningful use of an expression that in a novel situation there will be a standard of correctness for whether or not the term is applicable in this new instance. This depends precisely on an acceptance of grasp of ratification independent patterns of use. McDowell writes,

> Wright suggests (Wright 1980, pp.217-220) that the emergence of a consensus on whether, say, to call some newly encountered object "yellow" is subject to no norms. That is indeed how it seems if we allow ourselves to picture communal language in terms of

---

1. McDowell takes his argument to issue in a refutation of anti-realist conceptions of meaning and understanding. I think that he succeeds in refuting the anti-realist position as he sees it but that his restriction of the anti-realist to descriptions of behaviour which do not presuppose norms is, although suggested by some of Dummett's remarks and redolent of Wright's way of thinking, not well motivated. I have argued this postion in chapter 2.

> sub-"bedrock" resemblances in behaviour and phenomenology. But if we respect Wittgenstein's injunction not to dig below the ground, we must say that the community "goes right or wrong" ... according to whether the object in question is, or is not, *yellow*; and nothing can make its being yellow, or not, dependent on our ratification that that is how things are. (McDowell 1984, p.353)

The point here is that if we attempt to describe meaning in terms of behaviour characterized in norm-free terms then we will, in effect, be reduced to making inductive extrapolations in order to predict our reactions in any novel situation. Such judgements are always liable to be overturned by the actual behaviour of the community. But this means that there is no standard which the community must abide by in using an expression in a novel situation. The solution is to characterize our use of, say, the expression "yellow" in terms which presuppose norms by situating that use within a practice. The possibility of language can then be undermined by a dissolution of regularities in behaviour but, provided we are not tempted to explain meaning in these sub-"bedrock" terms, we allow that our understanding of an expresion can be appealed to as providing standards of correctness for its application. So it is right to call an object yellow just in case it is yellow.

Note that what is at issue here is not the informativeness of this claim -we are assuming we have given an account of what it is to understand "yellow"- but the objectivity of being yellow. The point is that an object is yellow iff it is, in fact, yellow irrespective of whether or not anyone actually calls it yellow. (This differs from the realist claim that an object determinately either is or is not yellow although is unrecognizable as being one or the other.)

It is apt to seem as if the question has, here, been begged. But the point was that we do not need to accept any account which is sanitized of norms. There is thus always a norm to which we may appeal. So a given use is correct if it satisfies the appropriate norm rather than if it elicits general communal consent. The distinction between realism and this form of anti-realism is, it seems to me, captured well in the following remark of McDowell's,

> The point about finitism is this. It recoils, rightly, from the mythology of the super rigid machinery - the patterns that extend of themselves, without limit, beyond any point we take them to. [This is realism.] But it equates this recoil with rejecting any conception of patterns that extend, without limit, beyond any such point. (McDowell 1984, p.353)

The first position which is rightly rejected is the realist position of believing in deteminate truth values which transcend recognition. The second is the acceptable intuitionist position of holding that we grasp verification transcendent truth conditions. The distinction was one I recalled towards the begining of this chapter. There I tried to argue that a finitism which rejected both of these positions is untenable. I hope now to have given some support to the intuitionistic position by showing that the anti-realist can and should accept the use the intuitionist needs to make of this notion of a variety of, what might be called, verification transcendent truth.

§10 The Revisionary Implications of Strict Finitism:

In this section I want briefly to question whether certain of Wittgenstein's considerations (about suveyability of proofs) should be taken to have revisionary implications. Whether or not Wittgenstein was a strict finitist in any relevant sense of the term, his ideas have greatly influenced the formulation of that position. Wittgenstein, at various stages in his *Remarks on the Philosophy of Mathematics* (1978), talks about the need for a proof to be surveyable. A proof is to give us certainty and must not only convince us that its conclusion is true it must also show us how it is true. Thus it must be absolutely perspicuous. A proof convinces us of the way things must be, it prepares us for adopting a criterion of correctness for the use of that procedure which is the subject of the proof. If it is to fulfil this normative role there can be no room for residual doubt over the correctness of the proof (which, of course, is not to say that we may not come to withdraw our assent to a proof: we can accept that we are fallible) since if our acceptance of a proof was consistent with harbouring a doubt about whether it was actually correct we would not, in circumstances where our operational procedure produced a result counter to the proof, be able to accuse the application of those procedures of having been mistaken. We would in the event of such a conflict have too many competing hypotheses.

Now Wittgenstein perceives this normative function of mathematics and develops an account of mathematical necessity from it. (See, for instance, Wittgenstein (1978), I §§75-105, III §§9-11, §§21-44, V §51, VI §23) In doing so he is not recommending the manner in which mathematics *should* function but describing the way it *does* function. For a proof to be treated as a proof it must be surveyable. Thus even if classical mathematics is to be criticized it is not to be faulted for using unsurveyable proofs. All the proofs of classical mathematics, in virtue of the fact that they can be treated as proofs, are surveyable. So no revisionary implications follow purely from acknowledging the need for proofs to be surveyable.

Wittgenstein often discusses the perspicuity or surveyability of proofs in the context of the logicist foundational enterprise (See Wittgenstein (1978), III §§12-20, §§45-64). The logicist attempts to justify our ordinary mathematical practice by showing that underlying every ordinarily accepted proof there is a proof relying only on sound principles provided by the logicist. Wittgenstein notes that the logicist's proof cannot provide the epistemological backing the logicist invests in it because the unsurveyability of the logicist's proof will in all cases mean that we shall accept the ordinary proof as giving the criterion of correctness for the construction of the logicist's proof. The logicist programme thus cannot provide epistemic access to mathematical truths since that is provided by the ordinary proof.

So Wittgenstein's emphasis on surveyability seems to have a criticism of foundational enterprises as its goal and to be lacking in revisionary import. It is not too difficult, however, to see that provided we take seriously the business of justifying our methods of proof a revisionary position does emerge. It is true that no putative proof, i.e., no construction which can be used as a proof, is unsurveyable but many proofs include steps which rely on there being determinate outcomes from applying an unsurveyable process. The thought then is that these proofs warrant criticism because they depend upon constructions which are, at most, mythical[1].

Denial of the justificatory programme draws the revisionary teeth of a finitist description of mathematics. Moreover Wright's deepest reason for rejecting intuitionism lies in what he discerns as a commitment to the strong objectivity of meaning entailed by that position. But if that is so where does the justificatory programme stand? Our criticism of non-finitistic practice centred on the use in proofs of methods of

---

1. Alexander George (George 1988 p.152 n.17) in criticizing Wright for not appreciating two different scopes of the intuitionistic use of "in principle" seems not to appreciate the point I am making here. George claims that Wright focuses on finite extensions of our actual recognitional abilities because he fails to realise that in his talk of the ability, in principle, to construct a proof the intuitionist need not be talking about extensions of our abilities but about our actual abilities to construct what is, in principle, a proof. But ultimately, if we are to justify this proof-in-principle we shall have to give some account of the legitimacy of considering the application of a process which we can only in principle carry out. The two positions George distingiushes seem to me to collapse into one another.

inference which depend on the existence of what, viewed finitistically, are mythical constructions, that is, putative constructions which a grasp of the meaning of our terms has not prepared us to accept as being fully determinate. From Wright's point of view, in understanding those terms we have made absolutely no commitment about how they are to be applied in these essentially new contexts. Our understanding makes no provision for what we must accept here so there is no meaning to which the proof must be faithful or which it can betray. The proof is thus unconstrained and can impose a decision on us about what the appropriate use should be. Thus any strict finitist position which follows Wright's motivation will not be revisionary. The point is that some intuitionistic practice will, if justified, appear to the finitist to import unjustifiable assumptions. But the underlying motive for that position questions the very enterprise of giving a justification. Abandoning that project means that considerations about the surveyability of proof have no revisionary import. So, for one thing, it is not clear why Wright, in particular, argues for a revisionary strict finitism. Also this position leaves it mysterious as to why we should accept a new proof and ignores the reasons, rehearsed earlier, for why we give a justification of methods of proof.

§11 Summary:

Intuitionism, in holding to the determinacy of truth values for effectively decidable sentences, does not suffer from the same defects as realism. The reasons for this are; first, that the intuitionist's use of provability is not analogous to the realist's use of bivalent truth or to the ideal verificationist's use of infinitary proofs; secondly, that the intuitionist can give an account of his understanding of vel which justifies use of Excluded Middle for decidable sentences. Allowing for this position means that the intuitionistic logic is not conservative relative to possession of canonical proofs but only relative to possession of a construction which yields a canonical proof. But since that notion is appropriately explained in terms of canonical proof there need not be a concern about whether the intuitionistic semantics contravenes molecularity requirements. What we do need to recognize is the complex relation between canonical and informal proofs within an intuitionistic semantics. This position is not at all untoward since the intuitionist must give some account of the place of informal demonstrations within mathematical practice as he rejects (for Gödelian reasons) the notion that arithmetic can be completely formalized. (See chapter 6 for more discussion of this.) The threat to intuitionism posed by a strict finitist position is based on the view that intuitionism makes illegitimate appeals to the character of our own understanding. But that argument hinges on a repudiation of the idea that there are ratification independent facts about correct meaning. However that sceptical attitude about meaning is not a position the anti-realist need have any sympathy with. Thus intuitionism has not been shown to be an unstable position.

CHAPTER FIVE: LOGIC AND SET THEORY

Logic and Set Theory

•

This chapter begins with an account of the intuitionistic logical constants. I consider the origin of undecidability in mathematics and the treatment of undecidable sentences in intuitionistic logic. This leads me to consider the account offered of our grasp of quantification over infinite domains. I draw out some links between that account and an account of our grasp of the set concept. I use these reflections to motivate a doubt about the definiteness of the unrestricted quantifiers in set theory. I attempt, finally, to gauge the feasibility of using that observation to motivate a coherent set theory which is rich enough to provide for an acceptable development of mathematical analysis.

The route towards a position in which it is claimed that we need to revise classical mathematics began with general considerations about the nature of meaning. Those considerations led to constraints on what we are to count as meaningful use of language: that we cannot credit ourselves with grasp of conditions governing the correct use of a term which are both determinate and lie radically outwith our recognitional capacities. Inferential practice, in particular, must accord with those constraints. So the project, in general, gains its revisionary bite by dictating which *logic* is applicable to a given region of discourse. We do not repudiate classical principles of inference in every instance. We do, of course, repudiate the realist justification of inference by appeal to a notion of truth which determinately obtains or fails to obtain irrespective of whether we can determine the matter and we also accept that use of classical principles of inference *unrestrictedly* entails a commitment to realism. Whether or not classical logic is applicable depends on the semantic account of the *content* of the sentences concerned. In this section I want to focus on reasons for thinking that in the mathematical case the content of some mathematical sentences exempts them from use in certain classically accepted inferences and thus for thinking that the practice of classical mathematics is in error. To that end we need to look at the sorts of sentences concerned and which inferences they fail.

Classical sentential logic is valid for decidable sentences. We first notice a realistic attitude to a given discourse when classical logic is applied to sentences for which we have (at present) no method guaranteeing a verdict on their truth value. So, to understand why certain mathematical sentences are precluded from use in all classically valid inferences we need to look at the source of undecidability in mathematics.

Since agreement in the use of an expression is agreement about its meaning an account of the meaning of an expression should detail conditions under which the sentence is correctly used or, in the case of declarative sentences, asserted, that is, we want an account of its of its assertibility conditions. A mathematical sentence is assertible just when we have a proof of it or when we have a demonstration that we are able (in some sense) to produce a proof of the sentence. So it is assertible just when it is prov*able*. Thus provability is taken as the central concept in the intuitionistic semantic theory.

Introduction of the notion of provability into the semantic theory must enable us to achieve the following: to explain conditions governing correct use; further, to explain these conditions in a way that does not depend on our grasp of *determinate* conditions which are outwith our capacity to recognize as holding; and, lastly, to account for these conditions in terms which do not presuppose a grasp of the the expression being explained. Our theory is thus reductive at least to the extent that it recognizes an obligation to give a substantial account of the conditions under which we hold a sentence to be true.

§1 The Meanings of the Intuitionistic Logical Constants:

Let us turn (as a prelude to discussing the intuitionistic account of undecidability) to the intuitionistic account of the meanings of the logical constants. The meaning stipulations can usefully be divided into two groups, the first given solely in terms of proofs and the second given in terms of operations which produce proofs when applied to proofs or members of the domain of discourse.

&, v, ∃

A proof of $A\&B$ is any construction that is a proof of A and of B.

A proof of $A \vee B$ is any construction of which it can be recognized that it yields a proof of A or of B.

A proof of $(\exists x)Fx$ is any construction that is a proof of $F\underline{n}$ for some n in the domain.

∀, →, ~

A proof of $(\forall x)Fx$ is any construction of which it can be recognized that applied to any n in the domain it yields a proof of $F\underline{n}$.

A proof of $A \rightarrow B$ is any construction of which it can be recognized that applied to any proof of A it yields a proof of B.

A proof of $\sim A$ is any construction of which it can be recognized that applied to a proof of A it yields a proof of $0=1$ (or of a contradiction).

Notice an obvious impredicativity introduced by the stipulation for the meaning of "→" (and of "~" which, in a sense, is derived from the meaning of "→", i.e., $\sim A$ is $A \rightarrow 0=1$) since in characterizing a proof of the conditional it implictly quantifies over proofs of the antecedent and since this domain has not been clearly circumscribed it may include proofs which involve the conditional whose proof conditions are being characterized. Thus the account we give of the meaning of the conditional will presuppose an understanding of the conditional and we seem immediately to have contravened the last of the three adequacy conditions I listed above. I shall not pursue this somewhat involved problem now but shall treat of it below in chapter 6.

A simpler point to note is that classical logic is validated for decidable sentences only. $\sim\sim A$ simply asserts that from a proof of $\sim A$ we can prove 0=1, i.e., from a proof that from a proof of A we can prove 0=1, we can prove 0=1. This is obviously weaker than being able to prove A itself since all it claims is that asserting that we cannot prove A leads to absurdity and that does not warrant us in asserting A unless we can assume that A is determinately either provable or refutable, i.e., that A is decidable. So the law of double negation elimination (DNE) and its equivalents fail for undecidable sentences.

§2 Undecidability:

It is now time to focus on the account of undecidability. In mathematics quantification over infinite domains is, at least in the first instance, responsible for the phenomenon of undecidability (conditionals may also be undecidable but only if either the antecedent or the consequent is undecidable and so ultimately the undecidability must issue from some other source).

Before considering the intuitionistic account of undecidability we should consider the nature of undecidability itself. A sentence is (now) undecidable if we have, at present, no means of definitely proving or refuting it. If we can show that a sentence can only be proved at the expense of a contradiction then that demonstration amounts to a proof of its negation. So the sentence is refuted. Thus we can never be in a position to assert that a sentence is, in principle, undecidable nor can we make the more general assertion that there are undecidable sentences since, on the intuitionistic interpretation of the existential quantifier, we have to be able to demonstrate an instance and that is just what we have shown that we are unable to do. We can, however, assert that not all sentences are decidable since this assertion simply demands that we draw a contradiction from the assertion that all sentences are decidable. The inability to assert the existence of an undecidable sentence does not contradict Gödel's Incompleteness Theorem since there is no bar to the idea that we are able to recognize that a given sentence and its formal negation are not provable *in a particular formal system.* The intuitionists also hold that the methods of proof which we can be brought to accept as valid are not circumscribable within the formal apparatus of a given system. Indeed the intuitionistically acceptable proof of Gödel's Incompleteness Theorem, it could be argued, demonstrates just this point.

§3 Quantification and Undecidability:

How does quantification over infinite domains give rise to undecidability? Although both classical and intuitionistic logicians will agree that quantification does give rise to undecidability their explanations for this will differ because of differences in the semantic theories they accept. The account given of what it is to grasp an undecidable sentence will differ from one case to the other (since the classicist, in contrast to the intuitionist, thinks that this grasp warrants attribution of determinate truth values to undecidable sentences) but both (obviously, since each accepts Gödel's Incompleteness results) will have to account for the fact that presently accepted methods of proof will not guarantee a decision procedure in every instance. Classically, the quantified sentence is interpreted truth functionally so (provided membership of the domain is both *a priori* and necessary - a plausible supposition in the mathematical case) it is logically equivalent to the infinite conjunction (for universal quantification) or disjunction (for existential quantification) formed by applying the predicate to each element in the domain. We must have this logical equivalence even though the classicist may hold that the content, whatever it is that we understand in grasping the quantified sentence, is not given in this way. (Witness Frege's remark (recorded by Dummett 1973, p.517) about the African chief who is unknown to him but who enters into the range of his quantified sentence about all men. Here Frege ridicules the idea that in using such a sentence he can be taken as intending an individual reference to that particular man. Clearly any classicist is going to be faced with daunting epistemological problems if he reckons that understanding a domain requires some sort of epistemic contact with each member of that domain.) Since we have no means of surveying an infinite totality we have no guaranteed decision procedure for sentences which quantify over infinite domains. So the quantified sentence can be undecidable despite the fact that it is interpreted as being logically equivalent to the infinitary truth function of decidable sentences which are each thus subject to bivalence.

Recall the intuitionistic interpretation of the quantifiers. Here the quantifiers are explained in terms of our ability to produce certain sorts of construction. Undecidability arises because, given any decidable predicate (i.e., one for which we have an effective means of determining the truth value of the sentence formed by applying the predicate to any object in the domain) we have no method guaranteeing that we can form an appropriate construction for the proof of the quantified sentence or that we can demonstrate the impossibility of such a construction. So, for the intuitionist undecidability arises not because we have a conception of a perfectly determinate domain which outruns our epistemic access but because the procedures by which we justify an assertion simply need not always issue in conclusive verdicts on truth value. The classical picture conjures up an image of our epistemic limitations in the face of our grasp of a concept (truth condition) which determinately "matches" or fails to "match" reality. The intuitionistic picture looks to the way we justify assertions within our practice and limits itself to outlining what we are justified in saying given that practice, whilst refusing to import metaphysical hypotheses in constructing justifications.

These remarks have been broad and, of necessity, vague, sufficient only to give an impression of the intuitionistic view. To get at the content of that position we must look much more closely at the intuitionistic explanation of quantification. Take the intuitionistic explanation of sentences of the form $(\forall x)Fx$, where F is a decidable predicate. To understand this sentence we must have the capacity to recognize of a construction that when it is applied to *any* member n in the domain it yields a proof of $F\underline{n}$. But this requires that we have some conception of any (or an arbitrary) member of the domain. So the characterization we have been given of what it is to understand a quantified sentence also quantifies, at least implicitly, over the same domain. How do we grasp this notion without having an intuitionistically unacceptable concept of the completed domain?

The query recalls Cantor's doctrine that each potential infinity presupposes an actual infinity. Grasp of a potential infinity is grasp of a notion of unlimited variability and grasp of an actual infinity is grasp of some completed infinite domain. Cantor denies that it makes sense "to speak of variability without speaking of variability over a completed domain" (Hallett 1984, p.25).

It might be thought that the threat of circularity in the intuitionist's use of "any" is no greater than his use of "and" in the explanation of conjunction. But we can very quickly see the evident disanalogy. In the case of "and" the conditions specified relate (at least in the case of decidable sentences, which are all we need consider for the moment) to conditions which are effectively recognizable. So in giving an account of the conditions in which a conjunction is assertible we can use "and" and rely on our ability to recognize the obtaining of just those conditions. That is, assuming that we grasp the proof conditions of each conjunct we do not presuppose a grasp of conjunction in requiring that someone be able to recognize that they possess proofs of both conjuncts. The problem with the quantified case is that we are in the process of characterizing conditions which are not guaranteed to be effectively decidable. This is revealed in the intuitionistic criticism of the classical explanation in terms of truth conditions which the intuitionist regards as valueless since circular. Take, for example, the classical explanation of universal quantification: $(\forall x)Fx$ is true just in case $F\underline{n}$ is true for all n in the domain. The explanation gives the truth condition of the universal sentence but in doing so itself quantifies over the very same domain. But wherein lies the difference between recognizing of a certain construction that applied to any n in the domain it yields a proof of $F\underline{n}$ and recognizing of the predicate F itself that applied to any n in the domain it yields the value true? The point here is that granting that in any instance a speaker can recognize a proof that F applies to some number n we seem to presuppose some grasp of quantification if we reckon that the speaker also grasps what it is for F to apply to any n.

The difference that the intuitionist wants to make good is this. He will not deny the classical explanation, after all the intuitionist's characterization of proof conditions is none other than a characterization of truth conditions, that is, of conditions in which we are justified in holding the sentence to be *true*. The difference lies in the nature of the notion of truth that the intuitionist regards himself as entitled to. So although accepting that we have a notion of truth which determinately obtains in each instance he will question whether this justifies assuming that this determines a unique truth value for the quantified sentence. The intuitionist holds that it is explanation of such a notion of truth which radically transcends verification which is problematic for the realist. The realist would explain this (if he offers any explanation at all) by saying that the truth value of each application of the predicate to an element in the domain is determinate (since the predicate is effectively decidable the intuitionist will not baulk at this) so the logical product or sum of these truth values gives a determinate truth value for the quantified sentence. So in the case of universal quantification the sentence is true if all the truth values of the instances are true and false otherwise. But, holding that one or other of these two possibilities must determinately obtain simply assumes the determinacy of truth value which the realist was supposed to be justifying (since we cannot effectively decide which possibility obtains).

A remark worth making at this point relates to the intuitionist's notion of truth. One attempts to justify explaining the notion of truth in terms of provability by appealing to conditions under which a mathematical sentence is regarded as assertible. But *provability* is not an effectively decidable property of a sentence, i.e., we cannot guarantee to be able to tell whether a given sentence is provable. The difference, as I've just noted, with the realist view is that truth as thus explained may transcend verification but cannot be assumed to do so determinately (i.e., *A* may be true although unverified but in the absence of a means of proof or refutation we cannot assume that *A* is true or false). So truth is constrained by our recognitional capacities, it is guaranteed a link with conditions which justify an assertion just because the proof relation is always decidable (i.e., we can guarantee to be able to recognize whether or not a given construction is a proof of a given sentence) and because we can only assert a sentence when we are in a position effectively to construct a proof. So we need never explain our concept of truth in terms of anything other than our recognitional capacities.

§4 Quantification, Objecthood and Predication:

We still, however, want more of an account of what it is for us to recognize of a construction that it acts appropriately on members of the domain. What is the nature of our concept of the domain if it is not a matter of surveying the domain? At this point in the discussion I consider both our understanding of domains of quantification and of classes and sets. I introduce classes and sets partly because of their intrinsic interest to mathematics but also because our concept of the domain of (pure) sets is, as a consequence of the set theoretic paradoxes, crucial.

The intuitionist will reject talk of bare existences, that is, talk of the existence of entities in the absence of considerations about how the entity might be presented to us. From a Fregean point of view this means that all objects must be, in some sense, nameable; objects are the referents of possible singular terms. This link between an ontological category and a syntactic one does not mean that objects come into being only when named nor does it involve a Meinongian explosion of our ontology, i.e., that an extra-linguistic object corresponds to each putative singular term. The first idealistic position is avoided (or, at least, delayed) because the existence of the object only depends upon its *possible* nameability so how much ontology relies on our lingiustic resources depends on how we gloss the notion of possiblity here. We do not necessarily arrive at the second position because the functioning of an expression as a genuine singular term is not purely a syntactic matter. Certainly we need to have syntactic criteria which determine putative singular terms (cf. Dummett (1973, pp.54-80); Wright (1983, pp.53-83); and Hale (1984)) but semantic properties are also decisive. In particular, the singular term must appear appropriately embedded in true sentences if the object corresponding to it is to be said to exist. It must also possess a genuine criterion of identity. That is, we must have a criterion by which we can recognize the bearer of the name as the *same* thing presented to us again and have the ability to distinguish the bearer of the name from other things. We can sum this up by saying that the understanding of a singular term requires that we can recognize its bearer as an instance of a sortal predicate where a sortal predicate is one which has an associated criterion of application (i.e., a determinate condition under which it is correct to recognize that the predicate applies to the object) and a criterion of identity under which we can re-identify and distinguish an object of that sort from other objects of that sort and can distinguish objects of that sort from others. So we gain the concept of some object only through an understanding of a range of objects, that is, through understanding the application of some sortal predicate. We cannot grasp a single object in isolation.

It is natural to suppose that grasp of a domain simply consists in grasp of a criterion of membership which is definitive of that domain: we understand the domain just when we know what condition an object must fulfil in order to be a member of the domain. (We need not be able effectively to determine membership since the criterion of membership need not always be decidable. For instance 1 might be a member of a domain only on condition that Fermat's Theorem is true.) So, it would seem that any intelligible predicate specifies a domain. Russell (Russell 1908, pp.72-3) notes that this is only true in the context of a pregiven range of significance for the predicate, that is, a range of objects for which the predicate is meaningful. But the point about the range of significance surely relates to our ability to form unique complements of the domain, i.e., to whether we can form the domain of just those objects which do not satisfy the criterion of membership and that this domain when united with the original domain yields the universe. It would seem we can, without previous restriction, consider just those objects of which the predicate is true. This description of what it is to grasp a domain does not depend upon us having a perfectly determinate grasp of the domain of all objects. That position seems to arise from considering reality as having a perfectly determinate constitution out of which we isolate just those objects possessing a certain property. But this is definitely not what we do. Rather we observe that each object must be capable of some linguistic presentation and that this presentation must always be definite enough to allow for predication. The predication can fail to have any sense, in which case the object fails the membership condition, or it develops a sentence capable of truth value, in which case the object is a candidate for membership (the candidacy is, of course, resolved by settling the truth value of the predication). The proposal is, at this stage, very generous in the sorts of domain *of objects* it is prepared to condone as intelligible. Indeed it allows sense to be made of statements about all objects provided we can credit ourselves with a grasp of what it is to be an object, provided, that is, we can confer determinate sense on the predicate "...is an object." If this were a consequence of the proposal it would not be entirely untoward since, as Dummett notes (Dummett 1973, p.530) we can meaningfully talk about all objects, e.g., all objects are identical with themselves. We do not immediately involve ourselves in paradox for two reasons. First, although we can form the putative singular term "the domain such that ..." we cannot always take it that this is a legitimate singular term. What is its criterion of identity? If this is given by coextensionality then we are considering domains to be sets. I shall presently consider this case. Secondly, we have only countenanced the development of domains of objects so all members of a totality must have a criterion of identity associated with them. So if we try to consider, say, the domain of predicates we shall have to consider what the criterion of identity is for predicates. If this is coextensionality then the case is again that of sets which I am about to consider.

Before embarking on that project I want to make a remark about Dummett's comment on sentences which unrestictedly quantify over all objects where he claims that we cannot take it that a sentence so formed always possesses a determinate truth value. It is somewhat mysterious as to what Dummett takes the relevance of this remark to be. On the one hand it is an uncontentious and general point about the intuitionistic treatment of undecidability. Why does Dummett recall it in the particular context of considering the paradoxes? It seems he intends to avoid the paradoxes in this fashion since he blames the classical assumption of determinate truth values for leading to contradiction, witness this remark,

> ...we cannot take the quantification over the totality of all objects as a sentence forming operation which will always generate a sentence with a determinate truth-value; in other words, interpret it classically as an infinitary conjunction or disjunction. If we attempt to do so, we shall be led into contradiction. (Dummett 1973, p.530))

But, if that is its intended import, it is simply false. Since proof by reductio is valid intuitionistically we can prove $\sim A \& \sim\sim A$ from $A \leftrightarrow \sim A$. Thus a paradox for a classical logician is one for the intuitionist as well. So if the intuitionist is to avoid that trap he must do so by motivating a departure from the naïve existence assumptions and not simply by avoiding classical use of LEM.

To summarize the position so far, I have tied grasp of a domain of quantification to grasp of predication. We grasp the general notion of quantification through competence in the practice of predication and we grasp particular domains by grasping (some of) those predicates which supply criteria of membership for the domain. This means that the problem of circularity which I tried to raise in the intuitionistic explanation of quantification vanishes. If a speaker grasps a predicate he grasps its general application to objects so he understands quantification at least to the extent of knowing what it is for *any* object to satisfy the predicate.

The final sentence of Dummet's chapter on identity in *Frege: Philosophy of Language* concludes, "Predication cannot be understood if we try always to construe it after the model of saying something about an object" (Dummett 1973, p.583). The point that Dummett is making arises from the view that a fundamental use of predicates consists in applying them to things picked out by use of demonstratives and, thus, to things for which we presuppose no particular well-defined criterion of identity. (Dummett is here arguing against Geach's view that when we talk of the sameness of objects we do so by presupposing a certain presentation of those objects as instances of specific sortals. According to Geach there is thus no absolute sameness relation but only a notion of identity relative to presentation.) But, despite Dummett's remark, it is still true to say that a full understanding of the predicate will involve grasp of what it is to apply the predicate to an object and this will presuppose, implicitly at least, grasp of an appropriate domain of application.

§5 The Need for Constraints on Set Formation:

A set is, for me, a domain for which it is legitimate to form the corresponding singular term by using the (set) abstraction operator. So essentially a set is a domain which can be treated as an object for which the criterion of identity is constituted by coextensionality, provided the set existence assumptions are legitimate i.e., (if $\hat{x}Fx$ and $\hat{x}Gx$ exist then) $\hat{x}Fx=\hat{x}Gx \equiv (\forall x)(Fx \leftrightarrow Gx)$. Now we know with the hindsight provided by the discovery of the paradoxes that naïve abstraction must be constrained in order to avoid paradox. There are only two sites at which any putative constraint can take hold: the definition of predicates, i.e., the range of intelligible predicates, or the treatment of the resulting domain as a set, i.e., as an object.

Consider the schema for naïve abstraction, $x \varepsilon \hat{z}Fz \equiv Fx$. Now under certain substitutions for the predicate F we develop a contradiction (or a theorem) when we take x to be or to be definable in terms which presuppose $\hat{z}Fz$. So, famously if we take F to be $x \not\varepsilon x$ substituting $\hat{z}Fz$ for x we get $\hat{z}Fz \varepsilon \hat{z}Fz \equiv \hat{z}Fz \not\varepsilon \hat{z}Fz$. The point is that in these cases we can either prove or refute the schema assuming only set theoretic properties of membership so that the predicate lacks a contentful criterion of application since so applied its criterion of application is either contradictory or vacuous. This observation seems to lend support to the idea that whatever constraints are necessary should be used in limiting what we count as intelligible definition of a predicate: certain putative predicates are simply nonsensical. But that reading of the situation would be over hasty. For one thing the predicate only seems to fail of meaningfulness when applied to the set or to a putative object whose existence is entailed by the existence of the set being defined. So it is still open to us to criticize the treatment of the domain defined by the predicate as a set. But if the predicate is meaningful it seems strange that we cannot subsume the domain so defined under the (ostensible) sortal concept of sethood and so treat it as an object. What assumption do we make in doing this? Traditionally it has been argued that what we assume is that such impredicatively defined objects are or can be members of the original domain; that in treating the domain as a set we allow the definition of, in some sense, essentially new objects and these we cannot take to be subject to the old, in some sense, superseded, condition of membership. Dummett sums the position up like this,

> If we have first succeeded in specifying a totality, then the use of individual variables ranging over that totality has a perfectly clear content, and we may employ it to form expressions for abstract objects: but we have no right to suppose that those objects must fall within the totality we originally specified. If, however, we attempt to characterize the totality by reference to the kind of expression which can stand for an element of the totality -as, in effect, Frege does- then, for our characterization of the totality to succeed, it must be supposed that the reference of each expression of that kind has been fixed independently of the specification of the totality. (*Frege Philosophy of Language* p.531)

So according to this way of thinking a set is illegitimate if it is supposed to include impredicatively defined objects or, alternatively, such objects are illegitimate if they are supposed to be members of the class in terms of which they are defined. I attempted a detailed discussion of a proposal of this sort in the first chapter where I considered Russell's advocacy of the Vicious Circle Principle. The question I want to investigate in the rest of this chapter is whether this repugnance for impredicative definitions is a consequence of an anti-realistically imposed revision of set theory.

§6 Anti-Realist versus Constructive Approaches to Set Theory:

Anti-realism is a global position: acceptance of the anti-realist's arguments commits one to anti-realism about all regions of discourse. The anti-realist does not attempt to construe the reality represented by our talk, say, of the past or future as having been *constructed* by us. Rather, the anti-realist simply refuses to assume that that reality has a determinate constitution independent of our means of coming to know the way it is constituted. Constructive or idealist imagery is thus not only not fundamental to the anti-realist's position, it is not a natural concomitant of that position either. Notwithstanding this, non-realism about mathematics is often enough identified with one or another version of constructivism. The practice of traditional intuitionistic mathematics is a paradigmatic attempt to view the subject matter of mathematics as consisting in the study of procedures which are idealized, but are never-the-less, implementable, in principle, by humans. Also, many attempts to motivate the axiom systems of set theory hinge on being able to substantiate a notion of priority of the members of a set over the set itself, i.e., the set is regarded as (somehow) being formed by "bringing together" its members which thus assume a certain priority over the set itself. This gives rise to (versions of) the iterative conception of set. (See for instance, Boolos (1983) and Wang (1983).)

Here, I want to investigate an approach to set theory which eschews this appeal to a constructive metaphysics. My reason for so doing is not primarily that I reject the traditional intuitionistic conception and the iterative conception as being invalid. More pertinently to my concern with anti-realism I simply do not see a constructive approach as a natural outcome of *anti-realist* arguments. The implications of applying those arguments to mathematics is precisely analogous to the application of those arguments to other regions of discourse, namely, to motivate a doubt about assuming that that aspect of reality is both determinate and epistemically transcendent. Anti-realism about mathematics should be a consisitent position which refrains from making that assumption in the practice of mathematics. The approach is, at least from the outset, attractive because many rejections of impredicative methods seem to rely on an underlying constructivism and because, conversely, impredicative methods pose a serious problem for constructive approaches. The hope thus is that by eschewing constructivism we may motivate a coherent set theory which admits impredicative methods.

There are two main approaches to explaining the concept of set, i.e., approaches which are not happy simply to take the set concept as primitive. The first considers the set as given by a law determining membership of that set. The second considers the set as a certain sort of combination of its members. I shall refer to these as the predicative (or intensional) and combinatorial approaches respectively.

The predicative approach is (badly) utilised in naive set theory. There, it is assumed that every property determines a set, i.e., the unrestricted comprehension "axiom" is assumed. Although this principle is, in view of Russell's paradox, a logical falsity it does respond to a powerful intuition that any definite property must categorize all objects into just those which possess the property and just those which lack it. (Note that even if we reject, following intuitionism, this appeal to LEM the intuition that a definite property defines a set, even if it is presently undecidable whether or not certain objects are members of the set, is still strong and, equally problematic. See above.)

The simple predicative approach of naive set theory must be rejected. But needing to reject the theory and finding the means to do so (i.e., diagnosing its precise errors) are two different matters. Clearly the treatment of a set as an object on a par with all others, combined with the view that all properties determine sets cannot be maintained consistently. One could try to argue either that only certain properties define sets or that the naive approach fails because it is insensitive to a "stratification" of the set theoretic universe. If one takes the latter view the combinatorial aspect of the intuition of set recommends itself since this is naturally associated with a sense in which the members of the set are "prior" to the set itself. (So, crudely a set cannot be a member of itself and there is no "static" universe of sets to form a set itself. The paradoxes can, in this way, be avoided.) Proponents of the former view might argue that a property defines a set provided that the set which putatively results is, in some sense, not too big. (It should be clearly noted that I am not suggesting that set theories can clearly be categorized as belonging to one or another of these two broad approaches. They cannot, partly because some attempts to motivate the "limitation of size" idea actively make use of the iterative conception which is a version of the combinatorial approach. (See Boolos (1971), Wang (1964); and Parsons (1975) and Hallett (1984) for criticism of these approaches.) Partly, this is also because some approaches combine predicative and combinatorial approaches, e.g., the intuitionist's use of both spreads and species.)

Most systems of set theory accept that the ancestral of the membership relation (i.e., ε^*, where $y \varepsilon^* x$ iff $\exists x_1 \ldots x_n (y \varepsilon x_1 \varepsilon \ldots \varepsilon x_n \varepsilon x)$ establishes an ordering relation on the domain of sets which is irreflexive, asymmetric and transitive. In addition, it is often required that we can have no infinite chains of membership of the form $\ldots x_2 \varepsilon x_1 \varepsilon x$, where the LHS can be extended indefinitely. Now, if this well-foundedness of sets is presumed to be a consequence of our intuitive concept of set, it would seem that it must be motivated by a combinatorial approach (making probable use of some constructive metaphor).

This should, however, not be too readily assumed since Russell's system in *PM* provides an exception. The system of *PM* adheres to a hierarchical ordering of sets (or classes) in which all sets are well-founded. However the motivation for this view is not based on a constructive view of sets. Russell there refrains from believing that classes exist so he certainly cannot believe that they are constructed. Instead Russell achieves this ordering of the universe of sets by taking a strictly intensional view of sets (sets are derived from propositional functions) whilst insisting that, as a result of the VCP, bound variables must be typed. This, of course, leads to the ramified hierarchy of propositional functions in terms of which sets are defined by making essential use of the axiom of reducibility. The lesson I want to draw from this is that there is a means of motivating a version of priority (or an ordering by the ancestral of membership relation) without resorting to a combinatorial approach and the attendant lure of a constructive metaphysics. The idea is that we can stick to a resolutely intensional characterization of set and look to the use of the *quantifiers* (or variables) in set theory to motivate the required ordering. To this extent the approach I want to explore might be called Russellian. The manifest problem with Russell's account is that his constraints on the variable lead to ramification with the consequent need to introduce the axiom of reducibility. When discussing Russell's work I noted that a lacuna in his programme arises because he fails to give an account of the sense of the variable. (I went on to note also that his involvement with ramification and reducibility show a concern with conditions of individuation governing certain sorts of item and that this concern should be founded on a theory of sense.) I propose to discuss an account which attempts to do justice to these considerations.

§7 Extensional Definiteness of Quantification:

The role of choice sequences in intuitionistic analysis is to provide, in effect, a liberalization of the domain of sets of natural numbers beyond that guaranteed by constructive functions. What I want to investigate here is whether we can provide, in an anti-realistically acceptable manner, a domain of functions or of sets of natural numbers which is a) sufficient for the needs of a theory of the continuum and, b) which does not make use of the combinatorial approach.

The issue concerns (as Dummett notes in his discussion of choice sequences) "the correct characterization of some particular domain of quantification and of the way in which its elements are given to us" (1977,p.451). I claim that, i) a domain of quantification is given by some definite condition or criterion of membership of the domain, and ii) that if quantification over the domain is to be regarded as extensionally definite (I shall amplify on this below) then the way the elements of the domain are given to us must make clear the criterion of identity (or, at least, of individuation) applicable to those elements.

The general point here is that having settled the characterization of the domain and the manner of presentation of its elements, the validity of statements formed by quantifying over that domain should flow purely from the meanings of the intuitionistic quantifiers and *not* from considerations to do with the ontological status of the elements of the domain. So, for instance, an assertion about whether or not certain choice and continuity principles apply to particular combinations of quantifiers, (such as $\forall x \exists \alpha$, where x and α range over, possibly distinct domains) is determined by the characterization of the respective domains.

The manner of presentation of elements of the domain is important for two reasons. First, it is only once we have determined this that the intuitionistic existential quantifier has a clear meaning. Asserting $\exists xFx$ intuitionistically demands that we be able (in some sense) to present some element, c, of the domain such that $F\underline{c}$ is true. The specification of the meaning of "$\exists$" only acquires a clear sense once we have specified what it is to be presented with an element of the domain. There is no direct argument from this aspect of quantification over a domain to other features of our grasp of the domain. In particular, we cannot argue directly from considerations about the manner of presentation of elements to a conclusion about the extensiveness of the domain. This point can be illustrated by reference to the theory of choice sequences. The following result holds (in most) developments of the theory (e.g., that of Kleene (1965) and of Troelstra (1969)), $\exists\alpha A(\alpha)\rightarrow\exists fA(\lambda"x.f(x))$, i.e., whenever we can assert that there is a choice sequence satisfying a given condition we must be able to supply a constructive function satisfying that condition. The result is a consequence of the intuitionistic meaning of "$\exists$" as applied to choice sequences: the only way we can be presented with a choice sequence is as a law like sequence, i.e, as a constructive function. But, of course there is no need to identify the domain of choice sequences with that of constructive functions (and, as Kleene notes (1965, p.47) to do so would be to falsify the fan theorem). Similarly, although we can only be presented with a function or set as a constructive function we should not be lulled into the belief that we only grasp a domain of constructive functions. (Here "constructive" means "effective".)

The second reason why the manner of presentation of entities in the domain is important concerns, what I have called, the extensional definiteness of quantification over the domain. Given a domain, we can, by quantifying over the domain, specify various sorts of item (e.g., elements of the domain, classes of elements in the domain, functions defined on the domain, etc.). These specifications are definite provided quantification over the domain is itself a definite operation.

Quantification over a domain is possible when that domain is determined by some (perhaps implicit) condition on members of the domain. However such quantification is not, I claim, guaranteed to be extensionally definite and so is not guaranteed to issue in definite specifications of other entities.

The extensional definiteness of the domain is not a function of the manner in which *particular* elements of the domain are presented. In this I am inclined to endorse Gödel's view that to accept a link between the existence of an individual and a particular mode of specifying that individual betokens an underlying constructivism (and here I am examining an approach which admits no need of applying a constructive point of view). However I do think that the extensional definiteness of a domain *is* a function of the *sorts* of individual which are candidates for being members of the domain. If the sort of individual in the domain has not been clearly determined we have no right to regard quantification over the domain as an extensionally definite operation since we have no conception of what further sorts of individual we might be brought to accept as being included in the domain, or, rather, the assumption that what possible sorts of individuals may be included in the domain is determinate is, itself, a form of realism. The image that this naturally conjures up is that we are not forced (contra realism) to "carve up" reality so as to see it as constituted of certain sorts of objects (i.e., reality has no discourse independent, metaphysical joints (with apologies to David Lewis)). But, having adopted a particular scheme, quantification is extensionally definite relative to that scheme. That is, the individuals concerned are not "brought into being" by our activity, they are there to be recognized by the means provided for by our scheme. It should be clear that I am not suggesting that quantification which is not made definite by an appeal to a given scheme is senseless or incoherent. Rather, I am saying that such quantification should be interpreted as being inherently vague or as making a certain open-ended commitment, i.e., as saying "Whatever scheme we adopt then...".

The proposal I want to explore is, in view of the above, the following. Quantification over a defined domain is *extensionally definite* (in the sense that entities specified by such quantification have been definitely specified) if and only if the sort of entity to be included in the domain is clearly circumscribed. The minimal constraint on what counts as a specification of the sort of entity is that a criterion of identity (or of individuation) should be supplied.

§8 An Attempt to Develop a Set Theory:

The task now is to apply this proposal in the development of a coherent, well-motivated set theory. I should say, in advance, that my conclusions about the feasibility of this project are tentative and, partially, negative. But that result is, in itself, not without interest. My first step is to diagnose a fault in the assumption of the extensional definiteness of unrestricted quantification over the universe of sets. The criterion of identity for sets is ostensibly provided by coextensionality. Coextensionality cannot however, on pain of circularity, provide us with a globally applicable criterion of identity. Consider the following statement of coextensionality,

$$\forall z(z\varepsilon x.\equiv.z\varepsilon y)\supset x=y$$

This supplies us with a definite criterion of identity only if the quantification used in the statement is extensionally definite. But that quantification can only be definite if the domain of quantification has a well-defined criterion of identity. The domain of quantification here is the universe of sets, so the quantification is definite only if coextensionality is, itself, definite. Thus there can be no non-circular justification of the definiteness of coextensionality as a global criterion of identity for sets: our grasp of the universe of sets must be treated as inherently vague, as perpetually subject to extension.

The use of coextensionality as a criterion of identity for sets is regarded universally as a distinctive feature of the set concept. So does the above argument simply discredit the notion of set? I do not think we should receive the argument as delivering this conclusion. The argument attacks the use of coextensionality as a single, globally applicable criterion of identity for sets. This does not mean that we may not be able to supply "ersatz" criteria of identity by restricting the range of quantification to extensionally definite domains. Moreover, we can use this requirement to motivate a suitable restriction on the concept of set (which, note, makes no use of a constructional metaphor). So we can, in broad terms, attempt to reconstruct or justify set theory by insisting, i) that the axioms governing set existence always allow for the application of a precise version of coextensionality, and ii) that comparatively more comprehensive versions of coextensionality extend conservatively more restricted versions of coextensionality.

As a move towards this programme I want to consider the emergence of certain constraints on set existence and how these constraints preclude the formation of paradoxical sets. We can only treat an object as a set if it has a definite criterion of identity in the form of (a version of) coextensionality. We then immediately have as a consequence of my argument that the universe of sets is not itself a set (i.e., there is no set of all sets). So Cantor's paradox is avoided. Also we can show that we cannot have sets for which the following hold: $x \varepsilon x$ or $x \varepsilon x_1 \varepsilon x_2 \varepsilon \ldots . x_n \varepsilon x$, for finite n. In the first case treatment of x as a set would depend on the definiteness of the criterion of identity for x, i.e., $\forall z(z \varepsilon x . \equiv . z \varepsilon y) \supset x = y$, where the range of quantification would perforce include all members of x. But then, were x to be a member of x, we would need a criterion of identity applicable to x for the quantification to be definite. So we cannot form a definite criterion of identity for x. Similarly in the second case we would simply have to have that the definiteness of a criterion of identity for x would presuppose that for x_n which would, in turn, presuppose that for x_{n-1}, etc., until, finally, the definiteness of the criterion of identity for x would presuppose itself. So we conclude, in particular, that the Burali-Forti and Russell paradoxes are avoided. It would be helpful to be able to conclude further that ε is a well-founded relation, i.e., that every non-void set has an ε least member, i.e., $\forall x[\exists w(w \varepsilon x).\supset.\exists y(y \varepsilon x \& \forall z(z \varepsilon y.\supset.z \not\varepsilon x))]$ or $\forall x[\exists w(w \varepsilon x).\supset.\exists y(y \varepsilon x \& \forall z(z \varepsilon x.\supset.z \not\varepsilon y))]$. However reasoning in the above manner all we can conclude is that we cannot have a set, x, for which every member of x shares a member with x (i.e., $\sim\exists x[\exists w(w \varepsilon x) \& \forall y(y \varepsilon x.\supset.\exists z(z \varepsilon y \& z \varepsilon x))]$: the argument to this effect is to draw a contradiction from the assumption that we had such a set, x. This assumption would only be valid if we could guarantee that x had a definite criterion of identity. But to guarantee that we should need to discount both the possibilities of a loop of membership and of a regress of membership, since either of these possibilities preclude formation of a definite criterion of identity. But if we can discount the possibility of a regress we must be able to say that for any member, y, of x we can only form finite chains of the form $x_1 \varepsilon x_2 \varepsilon \ldots y$. If we can discount the possibility of a loop of membership then all the x_i must be distinct, in particular, the chain cannot end with, say, $x_1 \varepsilon x_1$. Thus x_1 must contain no members, and *a fortiori* no members in common with x. This is not to say that x_1 is a member of x which does not share a member with x since we do not know that x_1 is a member of x. But, on pain of contradiction, not all of the x_i and y can be members of x which include only members of x_i, i.e., since the set is finite we can assert that at least one of the x_i or y is a member of x which includes no members of x. So we cannot have a specification of the set x. However we cannot conclude from this that we have an effective means for finding a minimal set. (This limitation will be of significance below.) There is, of course, also the problem of how we are to understand the quantifiers in the above statements.

This last problem hints at what is, perhaps, the most natural way of regarding the theory. First, we should recall that unrestricted quantifiers have not been accused of senselessness but only of extensional indefiniteness. So we could have a class theory in which unrestricted quantification (over sets) was used both in specifying classes (so we could have an unrestricted comprehension principle) and in specifying an analogue of identity (which I shall call equality) for classes. What we would be precluded from doing is treating classes as sets, that is, as objects capable of being members in other sets. We have to justify the assumption that a given class is a set by showing that that class has a precise defining condition and is subsumed within the scope of a definite criterion of identity. In this sense, the resulting theory would be reminiscent of aspects of von Neumann's set theory. The two approaches are distinguished by the fact that von Neumann justifies his treatment of a class as a set just when the cardinality of the class can be shown to be strictly less than that of the universe; he adopts a cardinal limitation of size theory (see Hallett (1984), chapter 8 for a pellucid discussion of these ideas). Here I have tried to justify treatment of a class as a set by constraints on definiteness drawn from considerations of the use of quantifiers in set theory.

The crude picture we are now presented with is of an extensionally indefinite theory awaiting precisification through a process of circumscribing domains of quantification. How are we to institute this process? Let me begin to draw out some of the difficulties that face us here by concentrating on an example that lies close to the original concern about analysis: with what grasp of the power set of the natural numbers can we credit ourselves?

From hence forward I shall be assuming without justification the axiom of infinity, i.e., that we have a domain $\underline{N}$ such that $\emptyset \varepsilon \underline{N}$ and if $x \varepsilon \underline{N}$ then $x \cup \{x\} \varepsilon \underline{N}$. I shall also be identifying the natural numbers with this set. What I am interested in exploring is whether, given that assumption, there is an anti-realistically acceptable means of forming a domain approaching that of the classical power set of the natural numbers.

If the domain of natural numbers is extensionally definite then we have a criterion of identity for elements in that domain. So, in particular, a domain formed by applying a definite property to this domain will select a subset of this domain which is extensionally definite. So we appear to have a justification of an instance of Zermelo's separation axiom,

$$\exists x \forall y (y \varepsilon x . \equiv . y \varepsilon \underline{N} \& \varphi)$$

We could take the variable y as ranging over the domain provided by $\underline{N}$. Then we could take $\exists \varphi \forall y (y \varepsilon x . \equiv . y \varepsilon \underline{N} \& \varphi)$ as setting up the condition of membership for the domain $P(\underline{N})$ which would have the criterion of identity $\forall z (z \varepsilon x . \equiv . z \varepsilon y) . \supset . x = y$, where the variable z ranges ove $\underline{N}$. That is, we could use the axiom of separation to set up the concept "subset-of-$\underline{N}$". We can then ask whether or not this results in a reinstatement of the classical power set by questioning whether we are able to define a subset of $\underline{N}$ by using properties which themselves involve quantification over $P(\underline{N})$ and also over arbitrary set theoretic domains. The first is necessary for the proof of Cantor's theorem: we assume that we have a one-one function, f, from $P(\underline{N})$ to $\underline{N}$, then we define $\alpha = \{x : \exists \beta (f(\beta) = x \& x \notin \beta)\}$. Then for $\alpha \varepsilon P(\underline{N})$, $f(\alpha)$ is defined and is a member of $\underline{N}$. Now ask is $f(\alpha)$ a member of α or not? If it is then, since f is one-one, we must have $f(\alpha) \notin \alpha$. So assume that $f(\alpha) \notin \alpha$. Then we have $f(\alpha) = f(\alpha) \& f(\alpha) \notin \alpha$, i.e., $\exists \beta (f(\beta) = f(\alpha) \& f(\alpha) \notin \alpha)$, i.e., $f(\alpha) \varepsilon \alpha$. Contradiction.

So no such f exists. Note that in order to define α we were required to quantify over $P(\underline{N})$, of which α is putatively a member. Further, if we are to credit ourselves with grasp of an *arbitrary* subset of the natural numbers we must allow as a definite property $(\alpha)\varphi(\alpha,\beta)$, where (α) stands for quantification over *any* domain of sets. Thus the question we face is, what is a definite property of the natural numbers?

It might seem that I have supplied an answer to this question: a property is definite just in case it is specified using only quantification that is extensionally definite. However, that answer simply leads us in a circle: a property is definite if it uses only extensionally definite quantification, conversely, we can only determine the extensional definiteness of the power class once we have settled the notion of definite property. Plainly, we cannot make any progress along this route.

A means of circumventing this difficulty might be to abstract from the manner of specifying subsets (and thus to ignore their mode of presentation) and consider the class defined by the membership condition given by,

$\forall x(x \in y . \supset . x \in \underline{N})$.

This, provided the quantification is extensionally definite specifies a definite set theoretic predicate. The thought now is that by restricting the domain of the variable x in the above to $\underline{N}$ (as we restricted the range of variation of z to $\underline{N}$ in forming the criterion of identity $\forall z(z \in x . \equiv . z \in y) . \supset . x = y)$ we can form the appropriate definite criterion of membership. But now it is apparent that the attempt in both these instances to make the criteria definite is muddled. The restricted forms of the criteria miscategorize infinitely many sets. Sets which include only elements of $\underline{N}$ and elements not included in the domain of quantification will be misclassified as subsets of $\underline{N}$ (e.g., $\underline{N} \cup \{\underline{N}\}$ is counted as a subset of $\underline{N}$). Similarly, infinitely many false judgements of identity will be made.

Clearly what we require is a means of "ignoring" sets which lie outwith the "intended" application of the precise membership and identity conditions. It is at just this point that one is tempted to draw on constructive principles to justify the claim that the awkward miscategorizations do not occur because the offending sets simply do not exist, or have not, at this stage, been formed. The problem with taking that tack (apart from any antecedent misgivings one may have about constructivisms) is to see how one combines this way of justifying the good behaviour of membership and identity conditions with the view that members of domains so defined can include impredicatively defined members; if the offending sets are not available to disturb the functioning of the membership and identity conditions then, surely, they are also not available to form specifications of members of the domain. As Hallett remarks on the feasibility of combining impredicative and constructive methods,

> No matter what powers of surveillance, or what ability to run through infinite collections in a finite time, are ascribed to a postulated constructing agent it seems to me that the constructing agent can never complete what we might call an impredicative process, i.e. can never construct a set (or number) via an impredicative definition. (1984, p. 236)

No constructive process can be postulated if a step in that process depends on the result of the process itself. So impredicative methods form an absolute bar to constructive methods. (It is partly because this reasoning seems to me very persuasive *and* because impredicative methods seem so inextricably involved with much of classical analysis that I think it is worth making this attempt to gauge the feasibility of a non-constructive, anti-realist set theory.)

One might suspect that the above shows that we cannot make sense of these envelopes of precise domains non-constructively, that these domains can only correspond to stages of development. So although the motive for restricting set theory would stem from an underlying argument about the definiteness of set theoretic quantifiers the only way of making sense of that notion of definiteness is through a constructive metaphysics.

A, seemingly, attractive response to this suspicion is to recall that unrestricted set theoretic quantification has not been shown to be incoherent but, merely, vague. Set theory could thus be developed along orthodox lines as an essentially vague theory. The theory *as a whole* is then justified by providing a precise interpretation of the domain of its quantifiers. In other words, we justify a set theory by providing a model for it. Two problems for this approach are immediately apparent. First, it is utterly mysterious how we are to be presented with the domains which are to constitute models for the theory. Secondly, we are concerned here with providing a justification of the (uncountable) power set of the natural numbers. It is evident that this approach cannot lead us to admit the exitence of an uncountable set since the Löwenheim-Skolem results show that any first order theory possessing a model possesses a countable model.

•

§8.1 Typing the Universe of Sets:

It seems to me that the only feasible attempt to implement the proposal must make sense of "stages" without appealing to an underlying constructivism. If a solution is to be found it must, in effect, make use of a means of a typing function on sets. In outline, the resulting project would consist in:

i) defining a type function on the universe of sets;

ii) defining extensionally definite domains of quantification by using the type function;

iii) justifying the axioms of a set theory.

I shall now give a sketch of how these steps might be carried out and shall then go on to raise certain difficulties for this plan of action.

•

i) *Defining Types:*

Recall that I am assuming that the axiom of infinity, i.e., the domain $\underline{N}$ is given. Then define T recursively as follows,

$T(x)=0$ if $x\varepsilon\underline{N}$

$T(x)=\mathrm{Sup}\{T(y):y\varepsilon x\}+1$ otherwise.

(Where 0 is $\emptyset$ and $x+1$ is $x\cup\{x\}$ and Sup is defined on sets of natural numbers such that $\mathrm{Sup}x=y$ where $y\varepsilon x$ and if $z\varepsilon x$ then $z\leq y$.) Note that if $x=z$ then $T(x)=\mathrm{Sup}\{T(y):y\varepsilon x\}+1=\mathrm{Sup}\{T(y):y\varepsilon z\}+1=T(z)$, provided only that we can assume that the type function assigns a unique value to each y such that $y\varepsilon x$ (or z). So, given this assumption the type function is well-defined. (Of course, we are assuming that a criterion of identity can be given, but not what form that criterion takes.) The definition of the type function thus depends of the validity of transfinite induction on ε. But, delaying that question for the moment, the definition still requires comment. A more general type function could be given in terms of the rank function: $\rho(x)=\mathrm{Sup}^{+}\{\rho(y):y\varepsilon x\}$, where Sup+ is the least strict upper bound. I have avoided this definition because it is blatantly impredicative and constitutes our definitional access to the ranks: $\rho(x)$ is defined as the element which is less than or equal to *all* elements possessing a certain property, which it, itself is supposed to possess. In contrast, the above definition simply supposes that Sup is an intelligible set theoretic function on a finite (even, if not decidable) set of natural numbers. (If the set is not demonstrably finite then we cannot assert that the function is defined for this argument. See below, where I discuss the need for "ceilings" on values of the type function.) The point here echoes Russell: impredicative methods are not suspect when they do not provide our sole access to the entity being defined. In the first definition of type we are assumed to have independent access to each element in the domain of natural numbers. So the type function does not provide a means of introducing a natural number or does not constitute our access to that number, it simply assigns (non-effectively) a natural number to any given set. The rank function, however, is not provided with an independent means of access to elements in its range, i.e., to the ranks.

I am assuming that T is defined on all sets and that T maps the universe of sets onto $\underline{N}$. The former assumption is even stronger in the light of the latter assumption: we can then define the predicate $S(x)$ for "x is a set" as $\exists y \varepsilon \underline{N}(T(x)=y)$, or better, given that the intuitionistic interpretation of $\exists$ would then require that we can effectively compute the type of a given set, denying $\exists y \varepsilon \underline{N}(T(x)=y)$ would be to deny the sethood of x. I am not able to justify this assumption but can perhaps allay some worries it might elicit. I assume, first, that we have an intuitive grasp of the domain of sets (gained through our grasp of predication) which suffices for the definition of a function on this domain. The assumption that this is legitimate further characterizes the nature of that domain. This should be acceptable provided that there is no tension between our intuitive concept and the legitimacy of the definition: the latter need not be seen as *arising out of* the former since what I am interested in is developing a structure which is rich enough to support "enough" mathematics and not in providing a complete analysis of our concept of set. The programme is thus to use the assumed intelligibility of the type function to define a domain of mathematical entities called "sets". The question then is, what mathematical structures will this domain support? We can begin asking that question by attempting to justify an axiomatic system for the domain. The rest of this chapter is an attempt to undertake just this.

Given the definition of types we have the following crucial entailment, $x \varepsilon y \rightarrow T(x) < T(y)$. provided $T(y) \neq 0$.

ii) *Defining Domains of Quantification:*

We want each domain to be extensionally definite and for the domains to form a cumulative hierarchy, i.e., we want $D_n \subseteq D_{n+1}$. Define the domains as,

$D_n = \{x: T(x) \leq n\}$

Then $D_0 = \underline{N}$ and $D_n \subseteq D_{n+1}$. Now each domain is given by a definite condition provided that the type function is definite. So each is extensionally definite if a criterion of identity can be provided for it. The constraints on a criterion of identity are that it conforms to a version of coextensionality and can be conservatively extended by increasing its range of quantification. *If* the domains are extensionally definite then we try to ensure satisfaction of this requirement as follows. Write $\forall x^n$ for $\forall x \varepsilon D_n$ then if we have $\forall x^n(x^n \varepsilon y.\equiv.x^n \varepsilon z).\supset.y=z$ as a criterion of identity for y,z, members of D_{n+1}, we want to show that this can be conservatively extended in the sense that for m>n,

$\forall x^n(x^n \varepsilon y.\equiv.x^n \varepsilon z).\supset.y=z$ iff $\forall x^m(x^m \varepsilon y.\equiv.x^m \varepsilon z).\supset.y=z.$

We try to prove this as follows,

(1) $\forall x^m(x^m \varepsilon y.\equiv.x^m \varepsilon z)$ iff

(2) $\forall x^m[T(x^m) \leq n.\supset.(x^m \varepsilon y.\equiv.x^m \varepsilon z)]\&\forall x^m[T(x^m)>n.\supset.(x^m \varepsilon y.\equiv.x^m \varepsilon z)]$ i.e., iff

(3) $\forall x^n(x^n \varepsilon y.\equiv.x^n \varepsilon z)$.

(4) Whence, $\forall x^n(x^n \varepsilon y.\equiv.x^n \varepsilon z).\supset.y=z$ iff $\forall x^m(x^m \varepsilon y.\equiv.x^m \varepsilon z).\supset.y=z.$

The transition between (2) and (3) is justified since, $T(y) \leq n+1$ and $T(z) \leq n+1$ so, if $T(x^m)>n$, then $x^m \varepsilon y$ and $x^m \varepsilon z$ are both false. So the biconditional, $x^m \varepsilon y.\equiv.x^m \varepsilon z$, is then always true, i.e., $\forall x^m[T(x^m)>n.\supset.(x^m \varepsilon y.\equiv.x^m \varepsilon z)]$ is true. Whilst the first conjunct in (2) is the same as $\forall x^n(x^n \varepsilon y.\equiv.x^n \varepsilon z)$.

The problematic transition is that between (1) and (2). It relies on the principle: $\forall x^m Fx^m$ iff $(\forall x^m \varepsilon D_n)Fx^m \& (\forall x^m \notin D_n)Fx^m$. The "only if" part of the biconditional is clear since a proof of the LHS supplies a proof of each conjunct. The "if" part is not clear. The RHS assures us that we have a construction which will take elements of D_n to an appropriate proof and a construction which will take elements of D_m not in D_n to an appropriate proof. But we do not have a construction which will take arbitrary elements of D_m to an appropriate proof, *unless* we can take membership in D_n to be decidable. So the proof depends on the computability of the type function[1]. (More on this below.)

(Note that the assumption here that the more inclusive domain is formed by one of the D_ms is not strictly necessary, we could repeat the argument for any more inclusive, definite domain. So we are not begging the question about whether appropriate D_ms exist.)

We now use induction to prove that the domains are extensionally definite. We assume that $\underline{N}$ is extensionally definite. Assume, for induction, that D_n is extensionally definite. Then we can quantify over D_n to form $\forall x^n(x^n \varepsilon y . \equiv . x^n \varepsilon z) . \supset . y = z$ which is the definite criterion of identity for D_{n+1}. Since, as noted above, each domain has a definite membership condition D_{n+1} must be an extensionally definite domain. So, by induction, each domain is extensionally definite. (Note that we could state this result as $\forall n(D_n$ is definite) since although the "n" seems schematic is is not, in fact, so: we can eliminate the schematic "n" by using the definition of D_n.)

So, assuming the computability of the type function, we can use it to define a cumulative hierarchy of extensionally definite domains of quantification.

<u>iii) *Justifying Axioms of Set Theory:*</u>

Axiom of Coextensionality:- From the above we have that if $T(z) \leq n$ and

---

1. It is true that the assumption of computability of $T(x^m)$ only spells out a condition in which *this* proof is intuitionistically valid. It has been pointed out to me (by Dr.T.Williamson) that an intuitionistically acceptable proof proceeding via a double induction on m and n can be given which assumes only that $T(x^m)$ is well-defined. So the restrictions on the set theory which I see as emerging through the constraint that the type function be computable may be more severe than they need be.

$T(y)\leq n$ then $\forall x^n(x^n\varepsilon y.\equiv.x^n\varepsilon z).\supset.y=z$. We also know that this can be conservatively extended. We can thus leave out the type index to get $\forall x(x\varepsilon y.\equiv.x\varepsilon z).\supset.y=z$, provided we ensure that the domain of quantificationis taken as being large enough. So, provided we can set an upper bound on the type of any given set, we can ensure that we have a definite statement of coextensionality which determines whether or not two given sets are or are not identical. But, since we have already had to assume that the type function is computable, this is not an *additional* assumption. The axiom of coextensionality is thus reinstated if (and only if) the type function is computable.

In justifying further axioms we can now assume that, provided an application of the axiom yields a set belonging to a definite domain then it has a well-defined criterion of identity. The general pattern for justifying the set existence axioms consists in, first, showing which domain the putative set would belong to *if* it were to exist and, secondly, then showing that it does exist by formulating a definite defining condition for it.

Axiom (Schema) of Separation:- $\forall x\exists y\forall z(z\varepsilon y.\equiv.z\varepsilon x.\varphi)$. Given any set, this axiom affirms the existence of a set comprising just those elements of the given set possessing a specified property. Given x, let $T(x)=n+1$ then, the elements of y must all be drawn from D_n. We can thus take y, if it exists, to be a member of D_{n+1}. To satisfy the existential condition we need to show that $\exists y^{n+1}\forall z(z\varepsilon y^{n+1}.\equiv.z\varepsilon x.\varphi)$. The only interpretation we can give to the unrestricted quantifier is in terms of a definite precisification which can be conservatively extended. Consider $\forall z^n(z^n\varepsilon y^{n+1}.\equiv.z^n\varepsilon x.\varphi)$. This is a definite set theoretic predicate defined on the domain D_{n+1}. If y_1 and y_2 (in D_{n+1}) both satisfy this predicate then they are the same (since $z\varepsilon y_1$ iff $z\varepsilon x.\varphi$, i.e., iff $z\varepsilon y_2$). So the predicate defines a unique member of D_{n+1}. Moreover the predicate can be conservatively extended in the sense that,

$\forall z^{n}(z^{n}\varepsilon y^{n+1}.\equiv.z^{n}\varepsilon x.\varphi)$ iff $\forall z^{m}(z^{m}\varepsilon y^{n+1}.\equiv.z^{m}\varepsilon x.\varphi)$, for m>n. (*)

This is because if $T(z^{m})>n$ then both $z^{m}\varepsilon y^{n+1}$ and $z^{m}\varepsilon x$ are false, so, $z^{m}\varepsilon y^{n+1}.\equiv.z^{m}\varepsilon x.\varphi$ is true. Whence, assuming the computability of types, we get (*) (in exactly the same manner as in the account of coextensionality). Thus given x, where $T(x)=n+1$, we have $\exists y^{n+1}\forall z(z\varepsilon y^{n+1}.\equiv.z\varepsilon x.\varphi)$. So, given any set we can (assuming we can compute its type) give a definite specification of its subset comprising the possessors of a given property, φ. In that sense the axiom of separation is justified.

Note that from the above we can only conclude that $T(y)\leq n+1$. There is no guarantee that we can effectively determine the actual type of y unless we can effectively determine the maximum type of elements of x satisfying φ, and, this, we cannot guarantee to be able to do. So we cannot assume that the type function is computable. I return to this source of tension presently.

Axiom of Union:- $\forall x\exists y\forall z(z\varepsilon y.\equiv.\exists w(z\varepsilon w.w\varepsilon x))$. This affirms the existence of a set containing all and only the members of the members of any given set. Take $T(x)=n+1$. If y exists it must be of type n. Since, if x is of type n+1 it must include a set, say, w, of type n, and none of any higher type. (That is, if we can affirm that x is of type n+1 then we must be able to affirm that it includes a set of type n and none of any higher type.) w must include a member of type n–1, which must be a member of y. So y is of type $\geq$n. Conversely, y cannot contain a set of type >n–1 since then x would contain a set of type >n. Therefore we need to show that $\exists y^{n}\forall z(z\varepsilon y^{n}.\equiv.\exists w(z\varepsilon w.w\varepsilon x))$. This is done, as above, by giving a definite set theoretic predicate for members of D_n and then showing that it can be conservatively extended. Note, first, that $\exists w^{n}(z\varepsilon w^{n}.w^{n}\varepsilon x)$ iff $\exists w^{m}(z\varepsilon w^{m}.w^{m}\varepsilon x)$, where m>n. From left to right this is easy since D_n is included in D_m. Conversely, if we have $\exists w^{m}(z\varepsilon w^{m}.w^{m}\varepsilon x)$ then we must have some $w\varepsilon D_m$ such that $z\varepsilon w$ and $w\varepsilon x$.

But if $w\varepsilon x$ then $T(w)\leq$n, i.e., $w\varepsilon D_n$. So $\exists w^n(z\varepsilon w^n.w^n\varepsilon x)$. We can thus ignore the index on w provided we ensure that it is greater than or equal to n. Consider $\forall z^{n-1}(z^{n-1}\varepsilon y^n.\equiv.\exists w(z^{n-1}\varepsilon w.w\varepsilon x))$. This is a definite set theoretic predicate of members of D_n. It is satisfied by at most one such set (since, assuming y_1, y_2 both satisfy it, $z\varepsilon y_1$ iff $\exists w(z\varepsilon w.w\varepsilon x)$, i.e., iff $z\varepsilon y_2$) . Finally, we need to show that, $\forall z^{n-1}(z^{n-1}\varepsilon y^n.\equiv.\exists w(z^{n-1}\varepsilon w.w\varepsilon x))$ iff $\forall z^m(z^m\varepsilon y^n.\equiv.\exists w(z^m\varepsilon w.w\varepsilon x))$, where m>n-1 (*).

This follows in the usual manner since if $T(z^m)>$n-1 then $z^m\varepsilon y^n$ is false and $z^m\varepsilon w$ is false (since $w\varepsilon x$ so $T(w)\leq$n, i.e., $T(w)\leq T(z^m)$). So $z^m\varepsilon y^n.\equiv.\exists w(z^m\varepsilon w.w\varepsilon x)$ is then true. Whilst if $T(z^m)\leq$n-1 then we get the LHS of (*). (*) therefore follows under the assumption of the computability of types. Thus given any set, if we can compute its type, we can specify its union set.

Axiom of Power Set:- $\forall x\exists y\forall z(z\varepsilon y.\equiv.\forall w(w\varepsilon z.\supset.w\varepsilon x))$ This affirms the existence of a set containing as members all and only the subsets of any given set. Note that if we are only interested in the power set of $\underline{N}$, $P(\underline{N})$, then we can get this from $\{x:T(x)=1\}$. The set theoretic predicate here is definite and the set is included in the domain D_1, so is a set of type 2. But now let us justify the general version. Take $T(x)$=n+1, then there is a set $w\varepsilon x$ such that $T(w)$=n. So $T(\{w\})$=n+1. $v\varepsilon\{w\}.\supset.v=w$ and $w\varepsilon x$ so $v\varepsilon\{w\}.\supset.v\varepsilon x$, for all v. Thus $\{w\}\varepsilon y$, if y exists. $T(y)$ is then $\geq$n+2. If $T(y)>$n+2 then there is a set z such that $z\varepsilon y$ and $T(z)>$n+1. So there is a v such that $v\varepsilon z$ and $T(v)>$n. But if $z\varepsilon y$ and $v\varepsilon z$ then $v\varepsilon x$. So $T(x)>$n+1, which contradicts the fact that $T(x)$=n+1. Thus $T(y)$=n+2, if y exists. We now need to show that $\exists y^{n+2}\forall z(z\varepsilon y^{n+2}.\equiv.\forall w(w\varepsilon z.\supset.w\varepsilon x))$. Consider $\forall z^{n+1}(z^{n+1}\varepsilon y^{n+2}.\equiv.\forall w^n(w^n\varepsilon z^{n+1}.\supset.w^n\varepsilon x))$. This is a definite set theoretic predicate of members on D_{n+2}. It holds of at most one member of D_{n+2}. Thus it defines a member of D_{n+2} provided that it can be conservatively extended. The conservative extension is possible but is not simple since we need to consider the relationship between the ranges of z and w. We have,

$\forall z^{n+1}(z^{n+1}\varepsilon y^{n+2}.\equiv.\forall w^{n}(w^{n}\varepsilon z^{n+1}.\supset.w^{n}\varepsilon x))$ iff

$\forall z^{m}(z^{m}\varepsilon y^{n+2}.\equiv.\forall w^{i}(w^{i}\varepsilon z^{m}.\supset.w^{i}\varepsilon x))$, where $m>n$ and $i\geq m-1$ (*)

Since if $T(z^{m})>n+1$ then $z^{m}\varepsilon y^{n+2}$ is false. Also there must then be some v such that $v\varepsilon z^{m}$ and $T(v)>n$. So $v\not\varepsilon x$. $v\varepsilon z^{m}.\supset.v\varepsilon x$ is then false. Now, if $i\geq m-1$ then v is in D_i (since v is a member of a member of D_m) so $\forall w^{i}(w^{i}\varepsilon z^{m}.\supset.w^{i}\varepsilon x)$ is false. So if $T(z^{m})>n+1$, $z^{m}\varepsilon y^{n+2}.\equiv.\forall w^{i}(w^{i}\varepsilon z^{m}.\supset.w^{i}\varepsilon x)$ is true. Whilst if $T(z^{m})\leq n+1$ then we get the LHS of (*). So (*) holds, assuming, of course, that types are computable. Provided then that we ensure that our ranges of quantification are large enough in relation to each other we can then extend our set theoretic predicate conservatively. So given a set we can definitely specify its power set provided that we can compute its type.

Axiom of Pairs:- $\forall x\forall y\exists z(x\varepsilon z.y\varepsilon z)$. This affirms the existence of a set containing any two given sets. This axiom is guaranteed by the cumulative nature of the domains. We know that if $n=\max\{T(x),T(y)\}$ then $x\varepsilon D_n$ and $y\varepsilon D_n$.

Axiom of Choice:- $\forall x\exists f(\forall y(y\subset x.\exists z(z\varepsilon y).\supset.f(y)\varepsilon y)$. This states that for any set there is a mapping which takes each non-empty subset of the given set onto one of its members. Quantificationally there are no problems with this axiom: if $T(x)=n+1$, the index of y can be taken to be greater than or equal to n+1 and the type of z will be less than or equal to n. Intuitionistically, the axiom is not controversial since if we know that the antecedent is satisfied we must have a construction of which it can be recognized that acting on any subset y of x produces a member z of y. But this is just to say that the appropriate f exists.

Axiom (Schema) of Replacement:-

$\forall x(\exists f\forall y\forall z(y\varepsilon x.z\varepsilon x.y=z:\supset.f(y)=f(z)).\supset.\exists w\forall y(y\varepsilon w.\equiv.\exists z(z\varepsilon x.f(z)=y))$. This asserts that the image of a set under a mapping, which is functional on the domain formed by that set, is itself a set. This axiom is falsified in the above development of a set theory since there is a functional mapping taking each n in $\underline{N}$ onto the domain D_n. So, given the sethood of the former, the sethood of the class of all domains would be guaranteed by replacement. That class is not however a set since it has no finite type. We could remedy this situation by repudiating the previous assumption that the universe of sets can be finitely typed. We cannot do without a type theory so the acceptability of the axiom of replacement would depend, at least, on our ability to give a non-question begging account of *general* definition by transfinite recursion, and that, in turn, would implicate the theory of transfinite ordinals. Although I cannot examine that programme here the worry would be that this version of the programme would presuppose mathematical structures which are as rich as those it attempts to justify. In particular, there is the concern I raised immediately after introducing the type function that impredicative methods would be presupposed in using the less restrictive rank function.

Axiom of Foundation:- $\forall x(\exists w(w\varepsilon x).\supset.\exists y(y\varepsilon x.\forall z(z\varepsilon y.\supset.z\notin x))$. This axiom asserts that every non-empty set contains an element with which it has a null intersection. I attempted to give a justification of this axiom earlier but noted that (even setting aside the question about the interpretation of the quantifiers) we could only justify the intuitionistically weaker principle that $\sim\exists x(\exists w(w\varepsilon x)\&\forall y(y\varepsilon x.\supset.\exists z(z\varepsilon y\&z\varepsilon x)))$.

If we now assume that recursive definition of types is legitimate then if the type of x is n (so x is non-empty) we can find a set $y\varepsilon x$, where the type of y can be at most n-1. If the intersection of x and y is non-empty then there must be a $z\varepsilon y$ with $z\varepsilon x$. z can be of type at most n-2. Again, we consider whether the intersection of z and x is null or not and then repeat the process. We must finally, i.e., after at most n+1 steps of this process, either discover a disjoint member of x or else an element of $\underline{N}$ which shares a member with x. Since the latter is finite we can test each of its members for whether it is a member of x not sharing any members with x (one of them must satisfy this condition). The problem with this description of what purports to be an effective procedure is that, at any stage, it may not be effectively decidable whether or not x and its member share a member. Only if we build in this condition can we ensure that the above describes an effective procedure. So, even under the assumption of typing, we can only justify the weaker principle since it only precludes the possibilities of an infinite regress in membership or of a failure in the ordering by (the ancestral of) membership. Both of these possibilities were also precluded in the earlier argument; they fail to deliver an effective means of determining an ε-minimal member of a non-void set.

Problems in the Account

There are a number of interrelated difficulties in the implementation of the programme sketched above. First, I have linked the use of coextensionality as a criterion of identity and the justification of the axioms (of choice, pairing, power set, union and separation) to the computability of types. The link has two facets. One facet is revealed by the need to prove that coextensionality and the axioms of set existence are capable of being conservatively extended and the other is manifested in the fact that the application of the criterion and axioms is seen to depend upon a verdict on type identity. Now, although the axioms of power set, union and pairing, can be taken to affirm the existence of sets of determinate types we cannot claim the same for separation. Separation only allows us to set a ceiling on the type of the set whose existence is affirmed in any application. The reason for this is twofold. a) The property used in defining a subset can, if the thrust of this approach is accepted, be specified using quantification and, in particular, impredicative quantification, provided that the quantification used in the specification is definite. We cannot guarantee that a property so defined will be a decidable property of members of a given set. So we cannot guarantee to be able to determine the maximum type of a member possessing the property. b) Even if the property is decidable we would need to quantify over the domain supplied by the original set in order to determine the maximum type of a possessor of the property. Although quantifying over that domain is guaranteed to be definite it may generate undecidablity. We cannot remedy this situation by jettisoning separation. To forego separation would be to admit that we cannot be sure that the intersection of any two sets is itself a set. Conversely, accepting the closure of sets under intersection would resurrect the problem since we cannot guarantee to be able to determine the maximum type of a set included in two given sets. So the justification of the axioms depends on an assumption of computability of types which is in tension with the axiom of separation.

This problem is, I think, pernicious. The inspiration for the introduction of types is the need to give a non-constructive means of stratifying sets. The purpose of that stratification is to allow explanation of precise criteria of identity by restricting the range of quantification to sets below a certain stratum. It then becomes a critical matter to show that the functioning of these criteria is unimpaired by excluding sets in higher strata. We attempt to demonstrate this by showing that the inclusion of such sets by increasing the range of quantification delivers criteria which are guaranteed to agree in judgements of truth value arrived at by using the more restricted criteria. It is just at this point that we wish to appeal to the properties of the type function. From a realist perspective this process seems successful and unobjectionable: the type function shows that, in restricting the range of quantification appropriately, we have included all sets "relevant" to determination of truth value. So nothing is disturbed in the passage to more inclusive domains, verdicts on truth value are preserved. From the anti-realist or intuitionist perspective we need to take account of our means of determining truth values. To guarantee the conservativeness of our criteria we should need to show that our epistemic position relative to the more inclusive domain was, in relevant respects, unchanged. But, as we saw, we cannot merely consider properties of the type function without considering how we may come to know those properties to obtain. So, for instance, we want to be able to assert that if the type of an entity is greater than n, say, it cannot be a member of the sets concerned (so the condition is trivially satisfied) whilst, if it is less than or equal to n it is subsumed within the original criterion. But to be able to assert this we must have some way of coming to know it, and this cannot be guaranteed (without demonstrating the computability of the type function). What the intuitionist needs, therefore, to justify the transition to the more inclusive criteria involves non-trivial progress in his epistemic position. The more inclusive criteria have thus not been demonstrated, inutionistically, to be conservative extensions of comparatively more restricted criteria.

The standing of general impredicative methods from a non-realist perspective now seems highly dubious. It is difficult, as I have noted, to reconcile such methods with any form of constructivism, whilst a non-constructive approach seems to have a powerful need for a type theory and that has only been vindicated on certain realist assumptions.

Secondly, a crucial aspect of this programme clearly concerns the legitimacy of the recursive definition of types. In traditional set theory one proves the legitimacy of recursive definition on a well-founded relation by induction. (Note that this proof, in general, requires the axiom of union in order to show that the domain of the function is the universe of sets. I have assumed that this is so and used that assumption to characterize the universe of sets. So I have not, effectively, assumed the axiom of union in justifying the axiom itself.) To justify the definition of types following this model we would need to show that ε is a well-founded relation (i.e., we would need the axiom of foundation) and we would need to justify induction on ε, i.e., if $\forall x[\forall y(y\varepsilon x.\supset.f(y)).\supset.f(x)]$ then $\forall xf(x)$. The justification of induction requires the least member principle, i.e., if $\exists xf(x)$ then there is an ε-minimal object for which f holds (if $\exists xf(x)$ then $\exists y(f(y).\forall x(x\varepsilon y.\supset.\sim f(x))$. Given the least member principle and the antecedent of the inductive hypothesis we then easily prove (classically) that induction holds by drawing a contradiction from the assumption that $\exists x{\sim}f(x)$, i.e., we prove ${\sim}\exists x{\sim}f(x)$, whence, classically, we have $\forall xf(x)$. However the proof fails intuitionistically since f is not, in general, decidable and the domain is not surveyable. A second problem concerns the least member principle which cannot be assumed for ε since that assumption is the same as assuming the axiom of foundation and that, as I have noted, cannot itself be intuitionistically justified.

It would seem that, given this, we cannot give an intuitionistic justification of induction on ε but must assume that it holds. That may not be an uncomfortable position for two reasons. First, consider the following classically equivalent statements,

(1) If $\exists xf(x)$ then $\exists x(f(x).\forall y(y\varepsilon x.\supset.\sim f(y))$.

(2) If ${\sim}\exists x(f(x).\forall y(y\varepsilon x.\supset.\sim f(y))$ then ${\sim}\exists xf(x)$.

(3) If $\forall x(\forall y(y\varepsilon x.\supset.\sim f(y)).\supset.\sim f(x))$ then $\forall x{\sim}f(x)$.

From (3) we derive classically the general induction schema by substituting ${\sim}f$ for f and using double negation elimination. (1) is a statement of a version of the least member principle which is equivalent to the axiom of foundation (which is thus classically equivalent to the induction schema). However intuitionistically the equivalences fail for, at least, the following reasons. We cannot make the transition from (2) to

(1) (which relies on the intuitionistically invalid law of contraposition: if $\sim A \rightarrow \sim B$ then $B \rightarrow A$). So we cannot assume that the axiom of foundation is intuitionistically interderivable with the induction schema. Thus any intuitionistic suspicion attaching to the former may not necessarily infect the latter.

Secondly, traditional intuitionistic attempts to justify transfinite induction are contentious. Most notably, Brouwer's proof of Bar Induction (which is demonstrably equivalent to the general principle of transfinite induction) relies on the notion of a fully analysed proof in assuming that we may fully circumscribe the sorts of inferences which might be involved in proving a statement of a certain form. Dummett recommends that rather than relying on such a proof we should assume the principle of Bar Induction as an axiom or axiom schema. (See Dummett (1977, pp.102-3).)

It would seem therefore that an anti-realistically acceptable set theory developed along the lines sketched above is not *incoherent*. The scope of that theory would be determined by having to ensure that methods of set formation guarantee that the type function remains decidable. So, at this point it is unclear what the scope of such a theory would be. A comparison of this approach with that of traditional intuitionistic use of choice sequences would be instructive.

§9 Summary:

This chapter began with an account of the preliminary explanations of the intuitionistic logical constants. I looked at the source of undecidability in mathematical discourse, namely, at quantification over infinite domains. I gave a description of our understanding of such domains in terms of a grasp of predication. This led me to consider possible approaches to set theory. I noted that there was no specifically anti-realist impulse towards a constructive (or combinatorial) approach to sets so that if such a position was adopted then it would have to result from a specific conception of *set*. My interest thus centered on the task of gauging the possibility and scope of a non-constructive anti-realist set theory. A major motive in that enterprize stems from the difficulty of reconciling the use of impredicative methods with a constructivism. The constraint that I used in motivating a set theory emerged from considering the extensional definiteness of quantification. For quantification to be extensionally definite, I argued, elements in the domain must possess a definite criterion of identity. So coextensionality was found to be circular if treated as a global criterion of identity. The project thus becomes one of circumscribing domains of quantification to allow for specification of precise criteria of identity and membership. That process requires a theory of types (which, in turn depends, upon the validity of transfinite induction). The promise of this programme was that impredicative methods could be reinstated provided that the appropriate domains of quantification could be shown to be extensionally definite. But a crucial step of this programme relies on the demonstration that criteria made precise by circumscription of domains can be conservatively extended. It was found that this is possible only if the type function is computable. That would bring about a severe restriction on the scope of the theory. It would thus seem that the promise to reinstate impredicative methods by taking this route remains, at present, unfulfilled.

CHAPTER SIX: CANONICAL PROOF

§1 A Problem in the Intuitionistic Explanation of the Logical Constants

§2 Complexity and Canonical Proof

§3 Normalization

§4 The Use of Normalized Proofs to Explain Canonical Proofs

§5 An Alternative Characterization of Canonical Proof

§6 Dummett's Supplementary Arguments for a Notion of Canonical Proof

§7 Summary

Canonical Proof

The intuitionistic semantical theory is an attempt to satisfy a philosophical hankering for a systematic account of the meaning of mathematical statements in terms of assertibility conditions, that is, in terms of conditions which we can perforce take ourselves to be capable of recognizing and which constitute a warrant for the assertion of the sentence. So, for each mathematical sentence the theory must give a systematic way of characterizing the *particular* recognitional capacity which, if exercised by a speaker, would justify us in ascribing to him an understanding of that sentence. In mathematics a sentence is assertible just when it is provable so the intuitionistic semantical theory takes provability as its central semantical concept. In other words, the intuitionistic semantical theory must give a specification for each sentence of what is to count as a proof of it. The specification given must be informative, it must not presuppose an understanding of the sentence being explained and the construction that the semantical theory sanctions as a correct proof must, of course, be one we are capable of recognizing as such. The first constraint is needed if the theory is to describe language in such a way that it is perspicuously learnable and the second to ensure that the theory does not lose sight of its point of describing meaning in terms of use. Thus the semantical theory can be criticized either for not issuing in informative specifications of the assertibility conditions of individual sentences or for achieving this only via the ascription of implausible recognitional capacities. A practice can be criticized if it can only be described by semantical theories which bear these faults. My purpose in calling to mind these general issues is simply to set in context the criticisms Dummett makes of the preliminary explanations of the meanings of the intuitionistic logical constants.

Recall the informal intuitionistic explanations of the meanings of the logical constants. There the aim was to give an account of the meanings of these constants by giving a recursive account of what constitutes a valid proof of a logically complex sentence, that is, to characterize what counts as a proof of a logically complex sentence in terms of a knowledge of what counts as a proof of those of its constituents which are combined under the dominant logical operator. In this way the meaning of each logical operator is, it is intended, captured. Plainly, whether or not these explanations do or do not ultimately confer coherent meanings on the logical operators they do set up a framework within which those meanings must be established. That is, if we take it that the stipulations must assume a notion of proof (so do not constitute coherent *explanations* of the logical operators) the form of the stipulations does constrain the acceptability of the notion of proof which can be assumed unless we are prepared to allow a radical dislocation between our inferential practice and the meanings of the logical constants. For instance, even if we accept classical proof procedures in the stipulations then the classical use of excluded middle unrestrictedly will not be a consequence of the meaning of disjunction as thus explained since a disjunction then will only be assertible when we have a proof of one or other disjunct and in many classical uses of LEM we do not have a *classical* proof of either disjunct.

§1 A Problem in the Intuitionistic Explanations of the Logical Constants:

What then are Dummett's reasons for suspecting the explanations are incoherent? Dummett's principle reason stems from the overtly impredicative character of the stipulations. He argues that if the notion of proof as used in the explanations is that of informal demonstration then the explanations are inevitably circular, the explanations fail to issue in informative specifications of the meanings of individual sentences of the language.

Consider the meaning stipulation relating to conditionals: a proof of $A \rightarrow B$ is any construction of which it can be recognized that applied to a proof of A it yields a proof of B. The problem, of course, is that in this stipulation we are required to quantify implicitly over proofs of the antecedent and if we take this totality to consist of all those informally acceptable demonstrations then we must allow it to include those informally acceptable demonstrations which proceed via the conditional itself (or via sentences of arbitrary degrees of complexity). But then our specification of the meaning of a conditional will have been given in terms which presuppose an understanding of the conditional itself (or of the entire language). The point is that under this interpretation of the explanations the notion of complexity becomes vacuous since it is impossible to impose a partial ordering on sentences. Or, to be more exact, a possible partial ordering using syntactic criteria has no semantic relevance. The need for a partial ordering of sentences by complexity results from the molecularity requirement. If we are to allow for our stepwise progression in competence we must show how this progression relates to a partial ordering of sentences. The manner in which we achieve this establishes the semantically relevant notion of complexity for sentences. So Dummett's problem becomes one of constraining the interpretation of the explanations to one allowing the imposition of a (semantically relevant) partial order on sentences. He requires no more and no less of a solution.

§2 Complexity and Canonical Proof:

Since our problem arose through interpreting the constructions used in the explanations as ordinary informal demonstrations it seems natural to look for a solution by distinguishing those canonical constructions which are used in the explanations from these informal demonstrations. So we must firstly change the stipulations so that they are framed in terms of canonical proofs and then add a stipulation which ensures that a partial order can now be imposed. Two forms of stipulation present themselves;

1. A canonical proof of A is a proof of A which does not involve sentences which are more complex than A.

2. A canonical proof of A is a proof of A which only involves sentences less complex than A.

The first problem with these definitions is the ambiguity in the notion of involvement in a proof. Is a sentence involved in a proof if it is actually proved in the course of a proof or does it simply need to be invoked as, say, an assumption? But, further, these definitions of canonical proof simply raise the question of how we give content to the notion of complexity. There are only three aspects of a sentence to which we can link the notion of complexity: its explanation; its proof; or its syntactical structure. So, we get the following three notions of complexity;

(a) A is less complex$_e$ than B if A must be understood prior to understanding B.

(b) A is less complex$_p$ than B if A must be involved in a proof of B.

(c) A is less complex$_s$ than B if A is a constituent of B.[1]

The question now is which gloss on the notion of complexity combined with which definition of canonical proof frees our explanations of conceptual circularity.

First, note a factor common to each notion of complexity. None establishes more than a partial order of sentences. Sentences may simply be unrelated by complexity. We can quickly see that this observation suffices to rule out, *ab initio*, the first (weaker) definition of canonical proof. Since, according to this definition, a proof may still be canonical if it includes a sentence unrelated by complexity to the sentence proved.[2]

---

1. An alternative defintion of complexity$_s$ suggests itself and I should explain why I do not consider it. Rather than define syntactic complexity in terms of the relation of being a constituent we could define it in terms of the number of appearances of logical operators in the sentence. In fact it makes little difference to my argument which definition is used since both cohere with the possible notions of complexity in the context of normalization.

2. This, of course, is not true if we take the definition of syntactical complexity as mentioned in the previous note since that notion does not set up a partial ordering. Here although every sentence is related to every other by complexity the relation is not antisymmetric so we only have a preordering ie. we can have sentences that are distinct yet are of equal complexity. We can rule out the first definition of canonical proof by showing that with sentences of equal complexity one may be included in a canonical proof of the other and vice versa.

So a conditional (A) which is unrelated by complexity to the antecedent of another conditional (B) will be involved in a canonical proof of the antecedent of B (simply by including redundant steps in its proof). Thus when we quantify over the canonical proofs of the antecedent we will presuppose an understanding of A. This in itself shows that we cannot set a bound on the complexity of sentences that need to be understood prior to understanding a given conditional. But it is also clear that we could find conditionals which were each unrelated by complexity to the antecedent of the other and in this way develop a conceptual circularity (simply by tacking the proof of the one conditional onto a proof of the antecedent of the other and vice versa). Thus, if a notion of canonical proof is to be achieved, it must emerge through a gloss on the notion of complexity as it is used in the second definition above.

This definition requires that any sentence sets a bound on the complexity of sentences involved in a proof of that sentence. So understanding the sentence will not, on this scheme of things, demand grasp of sentences of *arbitrary* complexity. But this still allows a circularity in the stipulations since if the notion of complexity we are using is not irreflexive then a canonical proof of a given sentence may involve the sentence itself. This, depending on one's interpretation, may or may not be harmful: in one sense all proofs involve the sentence being proved since that sentence is asserted as the conclusion of the proof. It would be well though, to insist in all cases that the notion of complexity be irreflexive. This need not beg any questions since we still have to settle the precise nature of involvement in a proof. Thus we use a notion of involvement in a proof which would lead to conceptual circularity were the notion of complexity not irreflexive. That is, involvement of a sentence, A, in a proof of B must lead to circularity when B is less complex than A or when B is A. So, since what we are trying to characterize is a proof recognitional ability, *A* is involved in a proof of *B* just when recognition of a canonical proof of *B* requires recognition of a canonical proof of *A*.

(a) will clearly not tolerate a non-irreflexive notion of complexity since if it were not irreflexive understanding of a given sentence would presuppose an understanding of that sentence. (c) can be made irreflexive by insisiting that the notion of constituent be restricted to proper constiuent. Finally, the irreflexivity of (b) will depend on the notion of involvement in a proof. Involvement in a proof is a relation between a sentence and a proof. So the notion of reflexivity does not apply to it directly. However we have insisted that we use a notion of involvement in a proof which is strong enough to prevent a sentence being involved in its own proof. So complexity$_p$ becomes irreflexive. It is important to realize that these are all adequacy constraints on our account. We still have work to do in explaining the possible accounts of complexity.

§2.1 Can Canonical Proof be Explained using Complexity$_e$?

Dummett's worry about the impredicativity in the stipulation about the conditional can be summed up in the above notation as a worry about whether every conditional is more complex$_e$ than itself. It might seem therefore that any proposal which set out to gloss our notion of complexity as complexity$_e$ would be bound to fail since it makes use of the very notion that is problematic. In other words, we want to satisfy ourselves that the explanations do allow for an order in understanding so it is illicit simply to appeal to the order of our understanding to evade this difficulty. But that crude observation passes over some of the subtlety of our position. The nature of this response is to accept the circularity revealed by Dummett and so to accept the need for a distinction between ordinary informal demonstrations and canonical proofs. Framing the explanations in terms of canonical proofs is, it is intended, sufficient to ban interpretations leading to circularity just because those interpretations are incoherent, i.e., they are illegitimate *as interpretations* since so framed they *cannot* be understood. The point is that we developed the incoherence by importing a notion of proof. But the explanations are intended to *set up* a notion of proof so they assume a naïvety in matters of proof on behalf of a recipient of the explanations. The character of our stipulation serves then to ensure that all proof procedures result from explicit implementations of the explanations, that is, we take seriously the recursive spirit of the explanations. Our effort to satisfy ourselves that the explanations are coherent now calls for us to scrutinize the form of the explanations themselves. If we do this we see that the stipulation at least has the effect of allowing us to rephrase the question we are facing. Instead of wondering whether each conditional is more complex$_e$ than itself we are now concerned with the question of whether a conditional can be less complex$_e$ than its antecedent. Clearly there is no general answer to this question except by appeal to something like the recursive spirit of the explanations. It might be felt that this aspect of the character of our understanding should be made more explicit. (I shall return below to consider why we might feel this need. For the moment I simply want to register that some, as yet vague search for explicitness pushes us to look further than this simple, perhaps simple-minded, solution. My aim in doing this is to try and clarify why Dummett is so drawn to a solution in terms of normalized proofs.)

§2.2 Can Canonical Proof be Explained using Complexity$_p$?

Our next proposal arises from coupling the second definition of canonical proof with the notion of complexity$_p$. This proposal has the odd consequence that (at least some) disjunctive sentences and existentially quantified sentences have no canonical proofs since it may be that no sentence must be involved in a proof of the disjunctive or existentially quantified sentence. So any conditional with such a sentence as its antecedent is trivially satisfied. Bracket, for the moment, though, this (serious) difficulty and let us consider whether the broad form of the proposal holds out any hope of solving the problem about the conditional. It does indeed seem promising since presumably it is always possible to prove the antecedent of a conditional without having to prove the conditional itself. So the proof involving the conditional will not be canonical and so will not have to be quantified over in a proof of the conditional. But how do we demonstrate this? We could do so either by "moving up" and considering the form of the stipulations themselves or by "moving down" and considering the structure of proofs. The problem is to show that complexity$_p$ establishes a partial order and this means showing that it is impossible for two statements each to be necessarily involved in the proof of the other. If we consider proofs as sets of sentences ordered by inclusion (that is, if we, very inadequately, refuse to read any internal structure into proofs and so treat the constructions merely as sets formed from the sentences used in the proof) then this ordering will not suffice to rule out this possibility, i.e., set inclusion properties alone will not preclude the possibilty that those sets corresponding to canonical proofs of *A* always include *B* and vice versa. Thus it is apt to seem that it is only by considering the third of our proposals based on the notion of complexity$_s$ that we get a fully explicit demonstration of the partial ordering of sentences. It is precisely at this point that the notion of normalized proof recommends itself as tailor made to suit our difficulty.

§2.3 Can Canonical Proof be Explained using Complexity$_s$?

The result which would succeed in rescuing the stipulations would be one which showed that there was no pair of sentences each necessarily involved in a proof of the other. The normalization theorem shows that each sentence has a proof which proceeds only via sentences which are of lesser complexity$_s$ than itself. So, by the normalization result, if we have two sentences, each of which is necessarily involved in a proof of the other, then we must have two sentences which are less complex$_s$ than each other, i.e., each must be a proper constituent of the other, which is impossible. So a normalization result would indeed solve our difficulties. But is not the normalization result much stronger than we need? Could there not be systems which satisfy the first constraint but for which normalization fails to hold?

Plainly, we cannot answer this until we give a more precise specification of the notion of being involved in a proof. Does this entail actually being proved during the course of the proof or does it merely entail being invoked as, say, an assumption? The use of normalized proofs seems to incline one towards the first interpretation since in a normalized proof each sentence is actually constructed in the process of constructing the conclusion. But if that is the correct interpretation then normalization is far stronger than we need since basic non-circularity requirements suffice for banning a situation where to prove one sentence we need to prove another and vice versa. Certain formalizations of classical logic will, of course, satisfy this requirement but do not possess normalization results. (See Prawitz 1965 pp.34-5). Recall that, since what we are attempting to characterize is a proof recognitional ability, mutual involvement is pernicious just when the involvement in each case implicates the proof recognitional ability for that sentence. So if A is involved in a proof of B it must be involved in such a way that it involves the ability to recognize a proof of A. There are only two ways in which this can come about. First, if A is proved en route to proving B and, secondly, if we use an operation on proofs of A, i.e., conditionalization (or negation). But this simply regenerates the original concern. So now the question of the relation of normalization and non-mutual involvement hinges on the question of whether there is a solution to the original problem which does not have recourse to normalization.

Consider, then, the syntactic version of the proposal. Obviously the problem about the meaning stipulation for conditionals is solved on this scheme of things since we cannot realistically doubt whether complexity$_S$ establishes a partial order. But our problems are not yet over since we still need to convince ourselves that proofs of the appropriate form actually do exist. A normalization theorem will satisfy us on just this point. Given my reconstruction of the debate a normalization theorem appears to be attractive because it avoids troublesome appeals to the character of our own understanding and to the nature of the explanations whose very coherence is at question. Also, normalization seems to exploit the most basic level of complexity, the very source of our compositional understanding which motivates the need for a molecular semantical theory. So, rather, than undermining the other notions of complexity we give content to those notions by focussing on the syntactical level.

•

§3 Normalization:

What is a normalization theorem? A normalization theorem is a result about the "geometry" of derivation within a given formal natural deduction system. A natural deduction system is one where the use of the logical constants is characterized by a set of introduction and elimination rules -one of each type of rule being associated with each constant- and in which the process of inference is examined in the construction of formal deductions based on introduced assumptions. So, for instance, the use of "&" is characterized by means of the introduction and elimination rules given respectively by,

&I) $\dfrac{A \quad B}{A\&B}$ &E) $\dfrac{A\&B}{A}$ $\dfrac{A\&B}{B}$

•

vI) $\dfrac{A}{A \vee B}$ $\dfrac{B}{A \vee B}$ vE) $\dfrac{A \vee B \quad \begin{matrix}(A)\\ C\end{matrix} \quad \begin{matrix}(B)\\ C\end{matrix}}{C}$

•

The normalization theorem tells us that it is always possible to rearrange a derivation to fit a particular pattern. To be more precise we show, in the normalization theorem, that any valid derivation can be rearranged to form another valid derivation of the same formula with the property that it contains no maximal formula, that is, contains no formula which results from the use of an introduction rule and is used as a premiss in an elimination rule. The derivation can thus be divided into two sections (either of which might be empty). The first consists solely of the use of elimination rules on the premisses of the derivation and on resulting formulae to construct progressively less complex$_S$ formulae. This is followed by a stage where only introduction rules are used to construct progressively more complex$_S$ formulae culminating in the conclusion of the derivation. Consider the following example,

$$\dfrac{\dfrac{\underline{\Sigma_1} \quad \underline{\Sigma_2}}{A \qquad B}}{\underline{A\&B}}$$

$$(A)$$

$$\Pi$$

This deduction is not in normal form because *A&B* is a maximal formula, it results from an application of &I and is immediately followed by a step of &E. The normalization theorem shows quite generally the possibility of making the obvious rearrangement of this proof to eliminate the maximal formula. In this case we get the following derivation of Π,

$$\underline{\Sigma_1}$$

$$(A)$$

$$\Pi$$

The introduction and elimination steps are seen to be redundant.(The theory has been described lucidly by Prawitz (1965).)

The points we need to take note of here are these. First, if a theorem is being derived then all the formulae used in the derivation are either less complex$_S$ than an axiom or less complex$_S$ than the conclusion. So normalized proofs can be partially ordered by the complexity$_S$ of the formula derived.

Secondly, elimination rules are inadmissable in the final inference of the derivation (unless applied to some formula less complex$_s$ than an axiom). (So Dummett's problem about the use of elimination rules in applying constructions deemed to be derivations of conditionals, negations and universally quantified formulae does not arise. I have not as yet considered this problem. I shall do so below but anticipate myself here simply to record that normalization does solve the problem.) Lastly, the normalization theorem is only a proof about the form of proofs within a specific formal system. What is the relation of these formal constructions to proofs in our intuitionistic mathematical theory?

§4 The Use of Normalized Proofs to Explain Canonical Proof:

Naturally Dummett wants to see a very close link here because the properties of normalized proofs are so well suited to solving both problems he has set himself. Here is Dummett on the relationship,

> ...normalized natural deduction proofs provide an exact analogy for what is required of canonical proofs if the intuitionistic explanations of the logical constants are to form a system free of conceptual circularity. Naturally, it does not follow from the normalization theorem for first order logic, that a similar theorem will hold good for any first order theory of intuitionistic mathematics. (Dummett 1977, p.396)

So Dummett wants the properties of normalized proofs to be mirrored exactly by canonical proofs. He also links the possibility of making sense of the notion of canonical proofs to the existence of a normalization theorem for a formalization of the theory considered. The canonical proof however cannot *be* the normalized proof since, if the canonical proof is to fulfil its meaning determining rôle it must be something for which a speaker's failure to recognize it as such, is to be taken as grounds for witholding or undermining an ascription of understanding. We cannot, however, tie a speaker's understanding to an ability to recognize a certain construction within some formal system. This would be wrong for a number of reasons. First, it fails to square with the nature of our practice where competence with ordinary informal proofs is taken as criterial and not simply good grounds for ascriptions of understanding. Also, the proposal is wildly implausible since recognition of a certain construction in a given formal system can only be constitutive of an understanding of the original theory if it is linked to an ability to recognize the system as a correct formalization of our original informal theory. It would be absurd to demand this level of self-reflectiveness about his practice from a speaker of the language.

We could try to relax the connection between canonical and normalized proofs but still maintain that, in some sense, the latter provide an analysis of the former. These difficulties, to do with the relation of the analysed proof and speakers' practice, would not, however, be obviated by the demonstration of a necessary ontological link (which is all an analysis could reveal) between certain informal proofs and formal normalized proofs, where the thought would be motivated by an intention to ascribe to speakers implicit knowledge of this link, since the basic difficulty arises from insulating speakers from the meaning determining aspects of their practice. Once we do this the link between the two sorts of *recognitional capacities* becomes, at most, contingent and, unless supported by an as yet utterly mysterious argument for identifying the two sorts of capacity, the demonstration of a necessary ontological link just misses the point. If the link between recognitional capacities which we actually exhibit and recognitional capacities constitutive of understanding is contingent we allow sense to be made of the absurd sceptical worry that we misunderstand our own language. So Dummett can only demand that the normalized proofs within a given formalization are a theorist's tool which point, in some natural way, to informal canonical proofs (where these are proofs which we actually might give and which would anyway undermine an ascription of understanding were they not recognized as such).

If one held that the purpose of our semantic theory was to provide a semantics for a formalization of any intuitionistic semantical theory then the worries I have just aired simply miss the point since then the adequacy of the semantical theory is a question raised *after* we have been given the appropriate formal theory. I take it that neither are we in this position nor should we be in this position. First, this cannot be the nature of our concern for, if it were, Dummett would not have raised his question about circularity arising in the intuitionistic explanations of the logical constants if the constructions used in those explanations are taken to be *informal* demonstrations. Secondly, the intuitionist holds that the meanings of his mathematical expressions are not to be identified with their use within any particular formal system. Dummett argues (in *The Philosophical Significance of Gödel's Theorem*) that Gödel's Incompleteness Theorem can be interpreted as a demonstration that the sense of mathematical statements should be given "in terms of the inherently vague notion of an intuitively acceptable proof, and not in terms of a proof within any formal system" (Dummett 1978, p.201). In *Elements of Intuitionism* Dummett clarifies this position by claiming that what has to be accepted is that any formalization of a mathematical theory supplies us with a means of expression which enables us to transcend the bounds of proof within that formalization. We can, however, maintain that, at any stage, the meaning of the logical constants is given relative to canonical proof as defined with respect to that formal system. In transcending the formal system the meanings of the logical constants (and, hence, of all mathematical sentences) change. The point is that the informal explanations must provide programmatic meanings, i.e., they must give the form of precise explanations of the meaning of the logical constants. I want to say that even if they are to fulfil this programmatic rôle the informal explanations should be seen to be coherent without this depending on our ability to come up with an appropriate formalization, since the informal explanations should give an account of our informal practice, albeit one which receives elucidation via formal examination, the meaning of the informal practice should not be made to depend on the formal elucidation. If we give meaning to the intuitionistic logical constants via a notion of canonical proof which is only given content relative to some formalization then that is precisely what we have done. In contrast, if we give the meanings of those constants relative to our present proof recognitional capacities then, although we might characterize those capacities more precisely relative to some formalization, i.e., show that those capacities have certain properties, provided we can recognize those explanations as coherent in the absence of such a formalization, we do not require the formalization simply to be sure of the meaningfulness of the practice.

Allowing, then, that there is a need for a formalization, the next and obvious question is to ask what makes for a correct formalization. In asking this question I hope to clarify further the limits of the role of formalization. Two essential desiderata immediately suggest themselves. First, each meaningful sentence of the theory should be associated with one well-formed formula of the formal system and vice-versa. This is to guarantee a morrorring of expressive power. Secondly, a sentence is to be provable just when its formal analogue is derivable. This is to guarantee a mirroring of inferential power. (Note there is no conflict here with Gödel's Incompleteness Theorem. Since, as Dummett notes, we can allow that informal methods of proof are indefinitely extensible and so not susceptible to complete formalization but also insist that, *at any stage*, the theory must be capable of being formalized. The extension of methods of proof then becomes linked to a change in our notion of canonical proof and so of the meanings of sentences in the theory.)

If we have satisfied these desiderata and can convince ourselves that we have done so then we can easily feel satisfied that canonical proofs with the appropriate properties do, in fact, exist provided only that a normalization theorem holds for the formal theory. Since then, for every theorem of the theory we have a derivation of its formal analogue. Moreover, by the normalization result we have a normalized derivation of the analogue. Since each formula of the derivation has an informal analogue we can "translate" the formal derivation into an informal construction. This construction will be an informal proof of the sentence since each inferential step sanctioned in the formal case by an explicitly given rule will be justified according to the meanings of the logical constants. (I shall review the status of this assumption below.) Since there is thus a close structural isomorphism between the two systems we can be sure that the informal proofs pointed out in this way will be proofs having the properties of normalized derivations and will thus be suited to fill the rôle of canonical proofs. This, I take it, must form the "bare bones" of Dummett's programmatic solution.

There is enough to question in this outline. How, for instance, are we to satisfy ourselves that the two desiderata are satisfied? There is a considerable assumption in the first desideratum in its supposition that we can completely capture the grammatical rules of our language in a recursive specification of allowable syntactical strings. Perhaps, however, it could be argued that such an assumption is presupposed by any search for a systematic theory of meaning. I shall anyway pass over this large question here. I mean to concentrate on how we demonstrate that the second desideratum has been fulfilled. We can only ensure this by giving an inductive proof on the complexity of sentences. That is, we would have to show that conditions warranting the use of introduction rules corresponded precisely to conditions which we take to govern the assertion of a sentence having the corresponding logical operator as dominant. We would, that is, have to ensure a match between the introduction rules and the meaning stipulations (or, at least, assume an understanding of the logical constants). We could then frame the elimination rules either by insisting on harmony with the introduction rules or by again resorting to the meaning stipulations. The choice does not much matter since the molecularity requirement means that we insist on harmony in both systems. So when we assumed above that the formal proof translated into an informal proof, that assumption was equivalent to holding that the second desideratum had been demonstrably fulfilled. The point must, however be granted: some appeal to the meaning stipulations is inevitable. I am concerned with two aspects of this. First, it seems to me to be doubtful whether such an appeal is legitimate if we are in the process of showing those stipulations to be *coherent.* Secondly, one of the recommendations of the syntactic approach was its promise to side-step such troublesome appeals to the meaning stipulations themselves. This promise has been shown to be empty. Does this make the project redundant?

Looking, now, to the accusation of circularity, it seems to be justified if we are presupposing an understanding of the logical constants. Since in this case a demonstration of the coherence of the explanations will depend on being able to demonstrate cetain properties about the notion of complexity$_S$. But demonstration of those properties depends on an understanding of the logical constants and so assumes the explanations are coherent. However, that is precisely what we set out to show. Of course, such a demonstration need not be entirely worthless since at root we might be convinced of the coherence of the explanations and this demonstration might, given that conviction, reveal why we are so convinced. But the demonstration does become impotent in the face of the sort of sceptical worries Dummett has raised about those explanations.

What is less clear, however, is whether we need to appeal to a full understanding of the logical constants or of the meaning stipulations. Is it not possible that we can implement the programme whilst relying only on an appreciation of the *form* of the stipulations? But it is at just this point that the programme is apt to seem redundant and for a much more direct exploitation of such an appreciation to seem more natural.

§5 An Alternative Characterization of Canonical Proof not Relying on Normalization:

So, looking, at least, to the form of the meaning stipulations is inevitable but is it coherent? I can see no reason why it should not be. No circularity is involved since we are not assuming that the explanations confer coherent meanings. Rather we are simply looking at those explanations in the light of a definition of canonical proof to see whether they are, thus interpreted, coherent. How else, indeed, could we attempt to exonerate the explanations?

When I discussed the possibility of a solution based on the notion of complexity$_e$ I noted that this approach allowed us to reduce the problem to one of whether a conditional can be less complex$_e$ than its antecedent (this rather than the question of whether a conditional is more complex$_e$ than itself). I also suggested that we might seek to make the recursive character of the explanations responsible for settling this question. I think that this line of response is roughly right and that it certainly captures our intuition that the explanations are understandable[1].

Compare the meaning stipulation relating to conditionals with that of, say, conjunction. We can give a proof of *A* which proceeds via *A&B* but we do not take this as grounds for suspecting that the explanation of "*&*" is circular. The reason why we do not is that the recursive character of the explanations means that we implicitly assume an understanding of *A* which can be arrived at independently of an understanding of *A&B*.

What then is the difference in the case of the stipulation relating to the conditional? Of course, the difficulty raised here is due to the implicit

---

1. The objection to basing the account of complexity on complexity$_s$ is not that this must be characterized formally but that use of complexity$_s$ requires that we *show* that the appropriate proofs exist. That demonstration demands, at least, an informal analogue of a normalization result. This could only be furnished, I claim, by arguing from the meaning stipulations to a set of rules governing the use of logical connectives. The essential point remains, i.e., in a demonstration of the coherence of the meaning stipulations we must make use of the form of the stipulations themselves.

quantification over proofs of the antecedent. Dummett's response is to seek to circumscribe the domain of quantification so as to ban the possibility of circularity. The proposal based on complexity$_e$ makes explicit use of the understanding we have of A arrived at independently of $A \to B$ to circumscribe the domain of quantification. Specifically, the fact that we can arrive at such an understanding means that there is no threat of conceptual circularity.

Now is this explicit appeal to such an understanding different in kind to the implicit assumption that is built into the whole recursive project? Well, let us look at the new stipulation which reads: a (canonical) proof of $A \to B$ is any construction of which it can be recognized that applied to a canonical proof of A (a proof of A involving only those sentences which must be understood prior to understanding A) it yields a proof of B. Since in any of the stipulations we assume that the sentences in terms of which the explanation is given are already understood the only additional assumption in this stipulation is that we can explicitly use this understanding to circumscribe a domain of sentences, i.e., we assume that if we understand A we know which sentences are required to gain such an understanding. That may be a controversial assumption and we may want more of an explanation about how we determine this (I shall consider this presently) but remember the proposal was mooted to solve a problem about a specific sort of circularity, its intention was to remedy a particular fault, i.e., we do not need to have a more explicit conception of canonical proof than is needed to ensure the coherence of the explanations and which can then be taken to emerge from the explanations themselves. We need now only question whether we can recognize whether or not a given construction satisfies the meaning stipulation for the conditional.

Faced with a putative proof of a conditional we have to determine whether or not it transforms canonical proofs of A into proofs of B. Now we may have no effective method for determining whether a given proof of A is or is not canonical but that is not quite the point. The use of the notion of canonical proof simply licences certain assumptions (which may themselves have to be argued for) about the sort of proof we need to consider. It would be unreasonable to expect that we could effectively circumscribe canonical proofs since the notion of a fully analysed proof is itself problematic and it is as well that we do not make the coherence of the intuitionistic meaning stipulations hang on making sense of that notion; which would be to make a large and dubious assumption about the nature of intuitionistic mathematics. Dummett's proposed solution in terms of normalized proofs embraces that option.

We do, however, want some account of the relation between proofs and canonical proofs, partly because we simply want some account of what makes for a valid proof *per se* and, partly, because although the meaning stipulation for the conditional is only given in terms of canonical proofs of the antecedent we also want it to be satisfied by all proofs of the antecedent. Normalized proofs provided a good model for this since the normalization theorem showed that any proof could be transformed into a normalized one so that if this transformation is followed by the transformation effected by the proof of the conditional we have a way of transforming *any* proof of the antecedent into a proof of the consequent. So there is a further project to be pursued but, I submit, that this enterprise begins with an acceptance that the intuitionistic meaning stipulations form a coherent, if strictly informal, system of explanation.

Characterizing canonical proof in terms of complexity$_p$ proved immediately problematic because disjunctive and existentential statements could, it turned out, be canonical proofless. The alternative image we get of canonical proofs in the light of this is that canonical proofs are minimal proofs in certain chains of proofs. That is, a proof of A is canonical if there is no proof of A properly included in it. We can take inclusion here to be given simply by set inclusion relative to sentences used in the proof. Note the reversal in our strategy here. Before we defined canonical proof in terms of complexity and then sought to explain the notion of complexity.

Here we have defined canonical proof directly and now have to show that it can be used to establish a partial order of sentences and more especially that it solves the problem about the conditional. So, we need to show that a proof of A proceeding via $A \rightarrow B$ cannot be minimal. Well, a proof of $A \rightarrow B$ is, by the meaning stipulations any construction which can be applied to a proof of A to yield a proof of B. Now either the proof of $A \rightarrow B$ is not used in the proof, in which case this redundant step can be eliminated so that the proof is not minimal, or it is used. But the only way we can use it other than in building up more complex sentences is to apply it to a proof of A. But then the proof must contain a sub-proof which is a proof of A. So again it is not minimal. This simple-minded consideration is, of course, only an informal version of a cut-elimination theorem. I think that we should accept this informal argument since, as I argued above, even the formal result will involve a similar appeal to the form of the meaning stipulations and the characterization of canonical proof does give us an effective method of determining, given any proof, whether or not it is canonical - simply check whether it contains a sub-proof. I think that this provides an adequate characterization of canonical proof.

The two notions of canonical proof, i.e., that based on complexity$_e$ and that on minimality of a proof sequence, are not quite the same. Since the first considers any proof canonical provided only that it uses sentences needed in the explanation of the conclusion: we can construct such proofs which are not minimal, e.g., we could have as a canonical proof of $A \vee B$ a proof of A which included as a redundant step a proof of B. But these slight differences are not really significant, we could, for instance, widen the notion of a canonical proof of C based on minimal proofs to include proofs which are either mininmal or which include only sentences occurring in other minimal proofs of C. (So if C were $A \vee B$ then a canonical proof of C would be a minimal proof of A (B) or a proof of A (B) which includes B (A), but no other sentences not included in a minimal proof of A or B. Clearly this cannot effect the established partial order since such proofs can only include sentences which are presupposed by *C*, and which thus cannot presuppose *C* themselves.) This would ensure coextensiveness of the two notions.

§6 Dummett's Supplementary Arguments for the Notion of Canonical Proof: Dummett has two other reasons for searching for a notion of canonical proof as distinct from that of informal demonstration. The first of these concerns the stipulations relating to the conditional, universal quantification and negation. Each of these stipulations characterizes the meaning of the logical constant in terms of an operation applied to proofs in the first and last cases and to arbitrary members of the domain in the second case. Dummett notices that if, in the application, we are entitled to use the elimination rule relating to that constant then the stipulations become vacuous: any contruction we choose to regard as being a proof of a conditional, say, will just in virtue of that *be* a proof of the conditional. This is because for the construction to be a proof of the conditional we must see it as transforming any proof of the antecedent into a proof of the consequent. If we allow modus ponens in the application of the construction then provided we see it as proving the conditional it satisfies the meaning stipulation since appending the construction to a proof of the antecedent followed by a step of modus ponens gives us a proof of the consequent. Dummett sums up the problem thus,

> The constraints on what constituted a proof of statements of these kinds would then all come from whatever intuitive prior notion of informal poof we were appealing to; the explanations of the logical constants would not, themselves, impose any constraints whatever, but would merely lay down conditions which are automatically satisfied, given certain elementary and indisputable properties of informal proofs.
>
> Obviously, however, this is not what is intended when the explanations of the logical constants are given; we are not appealing to

an already understood notion of proof. of which the validity of the elimination rules is partially constitutive, but laying down what is to count as a proof in such away that the validity of those rules follows as a consequence. (Dummett 1977, p. 393)

Why is not Dummett's second observation sufficient to allay the worry he at first expresses? An explanation need not be faulted simply because it makes certain assumptions about what constitutes a fit state for intelligent receipt of those explanations. Such assumptions are only problematic if they presuppose concepts which themselves require explanation. Patently this is not the case here where what is required of the pupil is that he is a complete novice in the practice of proof. So on this proviso the explanations succeed in inducting him into the practice if they are *correctly* received. The rub however lies in giving content to the notion of *correct* receipt of the explanations. Provided our pupil believes he has mastered the explanations those explanations will be powerless to disuade him. So in the event of a disagreement, say, the explanations cannot be resorted to to arbitrate. Thus we cannot make sense of the idea that the explanations serve to explain a particular practice.

We cannot excuse this state of affairs by considerations about "the limits of explanation"; that any explanation is subject to interpretation and so is apt to be misunderstood. The point here is that the terms of the explanation can be perfectly well understood and agreed upon and yet still they permit irresoluble disagreements to arise. But can we not use that consideration to solve the difficulty? Surely we can insist that a would-be practitioner must be prepared to justify *every* aspect of his practice as an explicit consequence of the explanations. The obvious repost to this is just to remind ourselves that if we allow the use of elimination rules then *any* construction viewed as a proof of a conditional will be a proof of the conditional by an explict implementation of the stipulation relating to conditionals. But how do we form the construction in the first place? It would be nice to be able to argue that either the construction must be formed in accordance with the meaning stipulations - so we have no reason for regarding it as non-standard - or it was formed in a manner not explicitly ratified by the meaning stipulations. In the latter case this will mean that we eventually uncover some use of a sentence involving another of the logical operators which is not sanctioned by the unambigous determination of its meaning. Tempting as this response is, it is not strong enough to resolve the difficulty since it allows that we can have as a proof of a conditional any construction with impeccable credentials but yet is not a proof of the conditional since it is a *bona fide* proof but only of some other sentence. So we have the absurdity that any proof can be treated as a proof of a conditional.

We cannot simply resolve the difficulty by, in some manner outlawing as candidates for proofs of the conditional, constructions which are proofs of other statements since the problem is one which affects all conditionals (and negations and universal quantifications). If we allow a construction to be a candidate for a proof of one conditional then it must be taken to be a candidate for proof status of any conditional. This means either that nothing is allowed to count as a proof of a conditional or that conditionals can all share the same proof. We cannot get round this "all or nothing" trap unless we have some way of distinguishing between how a construction is legitimately applied and how not. So we do need to look at the way the construction is applied.

Dummett's solution is, again, in terms of canonical proofs based on normalized deductions since, in the latter, use of elimination rules is excluded. This is a strange position for Dummett to assume. In the above extract from Dummett's appraisal of the problem he mentioned that "we are laying down what is to count as a proof in such a way that the validity of those [elimination] rules follows as a consequence." This observation is as true of introduction rules as it is of elimination rules. So relying on proofs characterized by introduction rules is equally to put the syntactic cart before the semantic horse.

The true source of the difficulty is the notion of application as used in the stipulations. The stipulations fail to inform us how a construction is to be applied so as to result in the appropriate proof. Indeed it might be just this point that Dummett has in mind, i.e., we informally "build into" the stipulation the corresponding introduction rule. So, a construction is a (canonical) proof of $A \to B$ if it can be recognized that when appended to an initial sequence of any (canonical) proof of A it yields a proof of B. And, a construction is a (canonical) proof of $(\forall x)A(x)$ if it can be recognized that a uniform substitution of n for a free variable in the construction yields a proof of $A\underline{n}$. Note that as a consequence of this latter stipulation we must assume that the induction schema provides us with a uniform operation applicable to members of the domain. We cannot then attempt to justify induction in terms of a finite number of applications of MPP since the number of applications will depend on the particular member of the domain and so the operation cannot be taken to be uniform.

Dummett is never very explicit about how he intends the notion of normalized proofs to be used. Since he links the possibility of a notion of canonical proof so closely with the possibility of formalization it seems clear that he intends the notion to be more than a loose model for how the meaning stipulations should be amended so as to give a coherent account of canonical proof. Indeed it often seems as if Dummett intends to give an independent account of the notion which can then be used in the explanations. (But in that case what purpose would the explanations serve?) If that is Dummett's project I think and think I have given good reason for holding that it is misguided. Whether, though, it is or is not his project I hope to have shown that a solution in terms of an informal amendment the stipulations is viable.

But my contentment is, perhaps, premature since Dummett argues that as part of an informal demonstration of a conditional it may be necessary to give a demonstration that the putative proof does indeed have the appropriate property of transforming proofs of the antecedent into proofs of the consequent. This point is still pertinent given my interpretation of the notion of canonical proof. Indeed, it is particularly so. I noted above that the notion of canonical proof allowed us to make certain assumptions about the form of proofs but that these assumptions would often have to be argued for. This is one case where an informal argument becomes integral to the functioning of the proof. Dummett envisages a case where it is simply not immediately recognizable that the construction satisfies the meaning stipulation and it requires an argument to convince ourselves of this. The problem now is that our efforts to confine the complexity of proofs which we need to consider via the notion of canonical proof threaten to come to nought since we have no way of circumscribing the complexity of these informal demonstrations. Dummett notes that,

> ...it is at just this point that any confines within which we seek to enclose the complexity of a proof of a given conclusion will burst. As a result, *the point of introducing the notion of canonical proof evaporates*; and we are once again faced with the danger that the intuitive explanations of the logical constants contain a vicious circle. (Dummett 1977, p.399) [my emphasis]

Prawitz simply dismisses this concern of Dummett as inviting a regress: we could equally well ask for a demonstration that the demonstration has the appropriate properties and so on ad infinitum. But that response totally misconstrues the nature of the problem. Dummett is pointing to a fact integral to the *actual* functioning of informal proofs. The problem of regress does not arise because informal proofs cannot be regressive: we cannot recognize a regressive structure as a proof.

The original point of introducing the notion of canonical proof was to circumscribe the complexity of proofs of a certain statement that we need to consider when quantifying over proofs of that statement. The reason for this was that without this proviso we are *forced* to quantify over proofs which include the conditional in the explanation of that conditional. So the notion was introduced to evade a definite circularity according to the explanations as they then stood. As Dummett notes we are here only faced with the threat of circularity so it is premature for him to conclude that the original purpose of introducing the notion of canonical proof has been compromised. To show that the two cases were similar it would have to be shown that just as we do have valid informal proofs of a sentence which use more complex sentences so we do use demonstrations, i.e., explanations of certain constructions, which use more complex sentences. In the first case we do have valid informal proofs which we do not want to outlaw as such yet which inevitably involve the explanations in circularity, i.e., we cannot but take the circularity that arises in this way as a reflection on the semantic theory. In the second case we only have the threat of a possible circle. The second case is thus much more like the stipulation relating to conjunction than it is to that of the conditional. After all, in a certain sense, since A can be proved via $A\&B$ there is some threat of circularity here.

The point against Prawitz, that we do not, in fact, involve ourselves in regressive explanations of the working of a conditional, also is a reason to wonder whether Dummett's worry need be taken to heart. If the question relates to the actual use we make of these explanations we need to be given some grounds for thinking that we do, in fact, use explanations which involve us in circularity. The point is that the explanations are used because they are psychologically appealing. If this is so then surely if they are to be effective then they cannot be circular. On the other hand if we dismiss this appeal to psychology why can we not simply treat the explanations as so much mathematically irrelevant but psychologically forceful accretion? We can allow circular explanations in the sense that these may have some heuristic value but then they are not treated as integral to the proof. So we do not need to have a method for banning in advance explanations importing a circularity. We can simply take the development of a circle as reflecting on the explanatory value of the putative explanation rather than on the coherence of the semantical theory. I am not suggesting that there is not an important programme of systematically demonstrating that the practice of intuitionistic mathematics is free of such circularity. Indeed I think that that is a sensible and profitable programme. But, first, it should not be thought that this programme is required as a demonstration that the meaning stipulations are coherent. And, secondly, the discovery of a circularity in our actual practice can be taken as reason to criticize, and so to revise our practice (of course the adequacy of the semantic theory will, in part, depend on the extent of revision thus demanded). I am urging a programmatic defence of the intuitionistic account of the meaning stipulations. The difference between Dummett, who also urges a programmatic solution to these difficulties, and myself, is that Dummett thinks that the coherence of the explanations depends on the success of the programme whereas I think that only the adequacy of the semantic theory is at stake.

Why is it then that Dummett sees such a forceful problem here? The answer to this lies, I believe, in Dummett's consistent contention that normalized deductions provide the means whereby we can establish a partial order within our mathematical language. Normalization establishes an explicit and rigid hierarchy. That certain proofs include elements which possibly transgress this ordering must be of vital concern to Dummett since the whole rationale for looking to normalization is challenged. So it is not that "the point of introducing the notion of canonical proof evaporates" but that the point of introducing canonical proof *as based on normalized deductions* evaporates.

Dummett has his own preferred solution. Once again it based on the notion of normalized deductions. The problem is that given a sentence C we no longer have a guaranteed method of circumscribing the complexity of proofs of C because those proofs may import arguments to the effect that the construction purporting to be a proof possesses the requisite properties. Dummett's solution is to replace *C* by *C** where *C** is the conditional whose consequent is *C* and whose antecedent is the conjunction of axioms of the form "*A*→there is a proof of *A* of such-and-such a kind" where *A* is a constituent of *C*. Then we consider canonical proofs of *C* as analogous to normalized proofs of *C**. The problem with this solution is that it is heavily programmatic and relies on making sense of the obscure notion of a fully analysed proof which has been questioned ever since Brouwer's use of the notion in the proof of the Bar Theorem. Dummett admits that "we have no clear idea of how to formulate such axioms, let alone how to find a set of general principles for generating them" (Dummett 1977, p.400) but, he goes on to say that, we have no reason for thinking such a task impossible. Perhaps not, but it is strange that the very coherence of the intuitionistic semantics should depend so thoroughly on a substantial assumption about the possible character of intuitionistic mathematics. Why should it be that the coherent description of a mathematical theory which adheres to general constraints drawn from considerations about meaning should depend on being able to give such a high level of explicit formalization? This result, if true, seems to me to be highly significant. I do not feel that it has, as yet, been established.

§7 Summary:

Although I concede that there is a useful role for normalization results for formalizations of intuitionistic mathematical theories I do not think that such a result is needed to establish, nor do I think that it can establish the coherence of the informal meaning stipulations. Those stipulations are required to guide any formalization, so the process of formalization cannot be used to show that they are coherent (it may however help to make certain aspects of our practice explicit). The coherence of the stipulations is presupposed in the formalization process. Conversely, I think that a notion of complexity is to be found by considering the form of the stipulations themselves. This less formal characterization of complexity has the welcome effect of not being firmly tied to any particular formalization. Thus when proof methods are extended so as to transcend the bounds of a given formalism we can see the resulting theory as still adhering to the informal stipulations. The development of the new theory can then be seen as an extension of the old theory, rather than simply as a new theory.

Conclusion

In this essay I have attempted, first, to motivate an anti-realist approach to some foundational issues in mathematics and then to argue that such a position coincides with a version of intuitionism. That conclusion is subject to a number of qualifications and reservations.

The coherence of adopting this approach depends on the ability to formulate succesfully and apply the anti-realist's semantic principles. Here I have offered no argument for anti-realism which would have any force for those who find anti-realism radically misguided. I have defended anti-realism from certain accusations that it is internally incoherent. Those accusations, derived from Haack and Wright, question the anti-realist's ability to concern himself with semantic justifications (of logic, in particular). I tried to give direct rebuttals both of Haack's attack and of aspects of Wright's attack. But Wright's position can only be comprehensively dismissed by showing that the interpretation of the Rule Following Considerations, upon which his position is based, is wrong. Rather than tackling that issue I have offered an alternative interpretation of the Rule Following Considerations which, I claim is consistent with anti-realism. So here there is a clear case of unfinished business.

The last three chapters deal with difficulties in the development of the anti-realist's positive programme for mathematics. Anti-realism is, however, a global position. So if it cannot be made to give a coherent account of *any* region of discourse its basic motivation will have been undercut. The anti-realist's positive programme for mathematics thus depends on his ability to generalize that account across language as a whole. For empirical discourse the anti-realist encounters the substantial problem of defeasibility.

My conclusions about the success of the positive programme for mathematics are, themselves, tentative. First, the implementation of that programme requires that we give clear content to the notions of canonical proof and of having a construction which issues in a canonical proof. The first must be explicable in terms which preclude the emergence of circularity in the intuitionistic stipulations of the meanings of the logical constants. I suggested that this may be possible by utilising the nature of our understanding of the form of those stipulations. The notion of having a construction which (recognizably) issues in a canonical proof is needed since I took this to be the notion with respect to which our logic applies conservatively. Both of these programmes need to be made more explicit before they can be accurately appraised.

Large questions still remain about the nature of an anti-realist set theory. I argued that the anti-realist should accept a constraint on his set theory which applies through the requirement that he furnish definite criteria of identity. That argument raises more general issues about our understanding of identity and of the logic of identity. I have not investigated these issue here. The power of the set theory which was finally sketched needs careful evaluation. Although the computability of the type function does not, in itself, preclude the possibility of some impredicative methods it seems clear that impredicative methods cannot, in general, be vindicated in this framework. So the question of whether we can, in this way, provide for an adequate theory of mathematical analysis remains to be answered.

The notion of extensional definiteness was used to constrain our use of quantification in giving definite specifications of items of various sorts. This reflection was then used to diagnose faults in the development of the set theoretic paradoxes. An interesting avenue of investigation lies in determinig whether and how this view should be extended to the intensional and semantic versions of the paradox. It would seem that these cases involve quantification over intensional items which, not being supplied with criteria of identity, is extensionally indefinite. So we cannot use such quantification to specify definite items. Although this promises to give a contradiction-free system, it also seems to be too restrictive: can we accept that all propositions which involve quantification over a domain of propositions are inherently vague? If not, when are such propositions definite?

VLADIMIR: I don't understand.

ESTRAGON: Use your intelligence, can't you?

Vladimir uses his intelligence.

VLADIMIR: (*finally*). I remain in the dark.

(Beckett, 1956, p.17)

Bibliography

Bar-Hillel, Y., and A.Fraenkel (1973) *Foundations of Set Theory*, North-Holland, Amsterdam.

Beckett, S. (1956) *Waiting for Godot*, Faber and Faber, London.

Benacerraf, P., and H.Putnam (eds.) (1983) *Philosophy of Mathematics: Selected Readings*, 2nd edition, Cambridge University Press, Cambridge.

Blackburn, S. and A.Code (1978) "The Power of Russell's Criticism of Frege: "On Denoting", pp.48-50", pp.65-77 in *Analysis*.

Boolos, G. (1971) "The Iterative Concept of Set", reprinted in Benacerraf and Putnam (eds.) (1983) pp.486-502.

____ (1987) "Saving Frege From Contradiction", pp.137-52 in *The Proceedings of the Aristotelian Society* 1986/7 Vol. LXXXVII

Boolos, G., and R.Jeffrey (1980) *Computability and Logic*, 2nd edition, Cambridge University Press, Cambridge.

Cartwright, R. (1987) *Philosophical Essays*, MIT Press, Cambridge, Mass.

Cocchiarella, N. (1980) "The Development of the Theory of Logical Types and the Notion of a Logical Subject in Russell's Early Philosophy", *Synthese* 45, pp.71-115.

____ (1989) "Russell's Theory of Logical Types and the Atomistic Hierarchy of Sentences", pp.40-62, in Savage and Anderson (eds.)

Copi, I. (1971) *The Theory of Logical Types*, Routledge and Kegan Paul, London.

Cummings, E.E. (1931) *Viva*, Liveright, New York.

Dummett, M.A.E. (1963) "The Philosophical Significance of Gödel's Theorem", reprinted in Dummett 1978a, pp.186-201.

____ (1973) *Frege: Philosophy of Language*, 2nd edition, Duckworth, London.

____ (1975a) "The Justification of Deduction", reprinted in Dummett 1978a, pp.290-318.

____ (1975b) "The Philosophical Basis of Intuitionistic Logic", reprinted in Dummett 1978a pp.215-247.

____ (1975c) "What is a Theory of Meaning? (1)", pp.97-138 in *Mind and Language*, edited by S. Guttenplan, Oxford.

____ (1976) "What is a Theory of Meaning? (2)", pp.67-137 in G.Evans and J.McDowell (eds.) (1976).

____ (1977) *Elements of Intuitionism*, Oxford University Press, Oxford.

____ (1978a) *Truth and Other Enigmas*, Duckworth, London.

____ (1978b) "What do I Know When I Know a Language?" lecture delivered at the Centenary Celebrations of Stockholm University, published by Universitas Regias Stockholmensis.

____ (1981) *The Interpretation of Frege's Philosophy*, Duckworth, London.

____ (1982) "Realism", pp. 55-112 in *Synthese* 52.

____ (1987) replies in Taylor (ed.) (1987).

____ (1991) *The Logical Basis of Metaphysics*, Duckworth, London.

Evans, G., and J.McDowell (eds.) (1976) *Truth and Meaning*, Clarendon Press, Oxford.

Frege, G. (1950) *The Foundations of Arithmetic*, trans. Austin, Blackwell, Oxford.

George, A. (1988) "The Conveyability of Intuitionism", pp.133-56 in *The Journal of Philosophical Logic 17.*

Gödel, K. (1944) "Russell's Mathematical Logic", reprinted in Benacerraf and Putnam (eds.) (1983), pp.447-69.

Goldfarb, W. (1989) "Russell's Reasons for Ramification", in Savage and Anderson (eds.)

Grattan-Guiness, I. (ed.) (1977) *Dear Russell-Dear Jourdain*, Duckworth, London.

____ (ed.) (1980) *From the Calculus to Set Theory 1630-1910*, Duckworth, London.

Griffin, N. (1980) "Russell on the Nature of Logic (1903-1913)", in *Synthese* 45, pp.117-88.

Griffiths, D.A. (1981) "A Reconsideration of Russell's Early Ontological Development", pp.145-52 in *The Philosophical Quarterly.*

Haack, S. (1982) "Dummett's Justification of Deduction", pp.216-39 in *Mind* 1982.

Hale, R. (1984) "Frege's Platonism", pp.40-56 in *Frege: Tradition and Influence*, edited by C.Wright, Blackwell, Oxford.

Hallett, M. (1984) *Cantorian Set Theory and Limitation of Size*, Clarendon Press, Oxford.

Heyting, A. (1956) *Intuitionism: An Introduction*, North-Holland, Amsterdam.

Hylton, P. (1980) "Russell's Substitutional Theory", in *Synthese* 45, pp.1-31.

____ (1990) *Russell, Idealism and the Emergence of Analytical Philosophy*, Oxford University Press.

Kielkopf, C.F. (1970) *Strict Finitism*, Mouton.

Kleene, S.C. and R.E.Vesley (1965) *The Foundations of Intuitionistic Mathematics*, North-Holland, Amsterdam.

Klenk, V.H. (1976) *Wittgenstein's Philosophy of Mathematics*, Nijhoff.

Langendoen, D.T. and P.M.Postal *The Vastness of Natural Languages*, Blackwell, Oxford.

Levy, A. (1979) *Basic Set Theory*, Springer-Verlag, Berlin.

Linsky, B. (1988) "Propositional Functions and Universals in *Principia Mathematica*", *Australasian Journal of Philosophy*, Vol. 66, pp.446-60.

Luntley, M. (1988) *Language, Logic and Experience*, Duckworth, London.

McDowell, J. (1976) "Truth conditions, bivalence and verificationism", pp. 42-66 in *Truth and Meaning*, G.Evans and J.McDowell (eds.) (1976).

____ (1981) "Anti-Realism and the Epistemology of Understanding", pp.225-49 in *Meaning and Understanding*, edited by H.Parret and J.Bouveresse (1981), De Gruyter, Berlin and New York.

____ (1984) "Wittgenstein on Following a Rule", pp. 325-63 in *Synthese* 58.

____ (1987) "In Defence of Modesty", in Taylor (ed.) (1987).

Moore, G.E. (1900) "Necessity", pp.176-93 in *Mind*.

Parsons, C. (1974) "Sets and Classes", pp.1-12 in *Noûs* No.8.

____ (1975) "What is the Iterative Concept of Set?" reprinted in Benacerraf and Putnam (eds.) (1983), pp.503-29.

____ (1983) "The Impredicativity of Induction", pp.132-53 in *How Many Questions? Essays in Honor of Sidney Morgenbesser*, L.Cauman, I.Levi, C.Parsons and R.Schwartz (eds.), Hackett.

Pears, D. (1989) "Russell's 1913 Theory of Knowledge Manuscript", in Savage and Anderson (eds.)

Prawitz, D. (1965) *Natural Deduction*, Almqvist and Wiksell, Stockholm.

____ (1977) "The Conflict between Classical and Intuitionistic Logic", pp.2-40 in *Theoria* 43.

____ (1978) "Intuitionistic Logic: A Philosophical Challenge", pp.1-10 in *Logic and Philosophy*, von Wright (ed), Nijhoff.

____ (1981) "Philosophical Aspects of Proof Theory", pp.235-78 in *Contemporary Philosophy Vol.1*, Floistad (ed.) (1981).

____ (1987) "Dummett on a Theory of Meaning and its Impact on Logic", pp.117-65, in B.Taylor (ed.).

Putnam, H. (1977) "Models and Reality", reprinted in *Realism and Reason* by H.Putnam, Cambridge University Press, Cambridge, (1983) pp.1-25.

Quine, W.V.O. (1936) "On the Axiom of Reducibility", pp.498-500 in *Mind* 45.

____ (1969) *Set theory and its Logic*, Harvard University Press, Boston.

Ramsey, F. (1978) *Foundations*, Routledge and Kegan Paul, London.

Russell, B.(1903) *The Principles of Mathematics*, Allen and Unwin, London.

____ (1905a) "On Denoting", reprinted in Russell (1956), pp.39-56.

____ (1905b) "On the Relation of Mathematics to Logic", reprinted in Russell (1973), pp.260-271.

____ (1906a) "On Some Difficulties in the Theory of Transfinite Numbers and Order Types", reprinted in Russell (1973), pp.135-64.

____ (1906b) "On the Substitutional Theory of Classes and Relations", reprinted in Russell (1973), pp.165-89.

____ (1906c) "On Insolubilia and their Solution by Symbolic Logic", reprinted in Russell (1973), pp.190- 214.

____ (1906d) "The Theory of Implication" pp.159-202 in *The American Journal of Mathematics*, Vol. 28.

____ (1908) "Mathematical logic as based on the Theory of Types", reprinted in Russell (1956), pp.57-102.

____ (1918) "The Philosophy of Logical Atomism", reprinted in Russell (1956), pp.175-282.

____ (1924) "Logical Atomism", reprinted in Russell (1956), pp.321-344.

____ (1956) *Logic and Knowledge*, editied by R.C.Marsh, Allen and Unwin, London.

____ (1973) *Essays in Analysis*, edited by D.Lackey, Allen and Unwin, London.

Savage, C. and C.Anderson (eds.) (1989) *Rereading Russell: Essays on Bertrand Russell's Metphysics and Epistemology*, Minnesota Studies in the Philosophy of Science: Vol. XII.

Shanker, S.G. (1987) *Wittgenstein and the Turning Point in the Philosophy of Mathematics*, Croom Helm, London and Sydney.

Stoll, R. (1961) *Set Theory and its Logic*, Dover, New York.

Taylor, B. (ed.) (1987) *Michael Dummett: Contributions to Philosophy*, Martinus Nijhoff, Dortrecht.

Tennant, N. (1987) *Anti-Realism and Logic*, Clarendon Press, Oxford.

Tiles, M. (1989) *The Philosophy of Set Theory*, Blackwell, Oxford.

Troelstra, A,S. (1969) *Principles of Intuitionism*, Springer-Verlag, Berlin.

Wang, H. (1964) *A Survey of Mathematical Logic*, Peking Science Press.

____ (1974) "The Concept of Set" reprinted in Benacerraf and Putnam (1983).

Wright, C.J.G. (1980) *Wittgenstein on the Foundations of Mathematics*, Duckworth, London.

____ (1981) "Dummett and Revisionism", reprinted in Taylor (ed.) 1987 pp.1–30.

____ (1982) "Strict Finitism", pp. 203–82 in *Synthese 51.*

____ (1983) *Frege's Conception of Numbers as Objects*, Aberdeen University Press.

____ (1985) "Skolem and the Skeptic", pp. 117–37 in *Proceedings of the Aristotelean Society*, Supplementary Volume.

____ (1986) *Realism, Meaning and Truth*, Blackwell, London.

Wittgenstein, L. (1961) *Notebooks 1914–1916*, Blackwell, London.

____ (1958) *Philosophical Investigations*, 2nd edition, trans. Anscombe, Blackwell, Oxford.

____ (1978) *Remarks on The Foundations of Mathematics*, 3rd edition, trans. Anscombe, Blackwell, Oxford.